

ABSTRACT

Title of Dissertation: COMPETENCE AND PERFORMANCE
IN THE DEVELOPMENT OF PRINCIPLE C

Megan M. Sutton
Doctor of Philosophy, 2015

Dissertation directed by: Professor Jeffrey Lidz
Department of Linguistics

In order to understand the nature of a given linguistic phenomena in the adult grammar, language acquisition research explores how children's competence with respect to such a phenomena develops. However, diagnosing competence can be challenging because it is not directly observable. Researchers only have access to performance, which is mediated by additional factors and is not a direct reflection of competence. In this dissertation, I explore a case study of children's early syntactic knowledge. My in-depth analysis of Principle C at 30 months provides novel insights into diagnostics for underlying competence by utilizing two distinct methods of analysis. The first analysis explores alternative mechanisms that have been proposed to account for early Principle C effects. By comparing across multiple linguistic contexts, I show that Principle C knowledge is the only mechanism which can account for all observed performance. The second analysis explores the deployment processes that are required to implement competence in performance. I present a novel analytic approach to identifying underlying knowledge which utilizes independent measures of these deployment processes. I show that

individual differences in syntactic processing predict individual differences in interpretation, implicating syntactic processing in Principle C performance at 30 months. Together, these findings extend our knowledge of the developmental pattern that characterizes Principle C, which can contribute to debates about the origin of this constraint as part of the grammar. This research provides new depth to investigations of children's early syntactic knowledge by highlighting new methods for diagnosing competence from observed performance.

COMPETENCE AND PERFORMANCE IN THE DEVELOPMENT OF PRINCIPLE C

By

Megan M. Sutton

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor Jeffrey Lidz, Chair
Professor Robert De Keyser
Professor Naomi Feldman
Professor Valentine Hacquard
Professor Colin Phillips

© Copyright by
Megan M. Sutton
2015

Dedication

To my amazing parents
For the road trips, care packages, and voicemails
For being there for me from 500 miles away

Acknowledgements

My graduate experience has felt very much like eating a habanero chili pepper on a dare.

You know that plenty of people have eaten these peppers; some people even do it more than once, so you think it can't possibly be *that* bad. Once you've begun, you regret your decision almost immediately. You can't imagine why you thought this would be a good idea. You realize that all the people you know who have done this are clearly sadists. You are encouraged by your peers to keep going, even though you are blinking back tears and your jaw is sore. You begin to understand humility on a deeper level than you previously thought possible. This task that seemed relatively simple at the outset seems to take an eternity; you start to forget what your life was like before you bit into this pepper.

Even after you've swallowed, you can tell that the burn of this experience is going to be with you for a long time. Maybe forever. But you quickly recognize that this is probably a good thing. Although it's a relatively small period of your life, you know you'll never quite be the same person you were before the pepper. Somehow, you are by turns exhausted and invigorated. To say you are a stronger person for having stomachached this adventure is an understatement. You begin to realize how much your experience has taught you about yourself. (You quietly wonder what it says about you that you are one of those sadists who might recommend this experience to another human being.) Even though you could probably find a lot of horrible things to say about what it was like to eat the pepper, retellings are always with a smile.

It is no exaggeration to say that I would never have been able to eat this pepper alone. The work presented in this dissertation has been shaped by so many that it would be near impossible to name them all here. Linguistics department faculty and students; teachers; classmates and colleagues; support staff; audiences for countless lab meetings, talks, and conference presentations—I can guarantee you've had an impact on this dissertation, whether you know it or not. Although I can't name you all, thank you from the bottom of my heart. You have made this an unforgettable experience.

Throughout the last five years, my advisor Jeff Lidz has been a beacon of unflinching support. He has been in equal measures a fantastic advocate and critic, and both of these have been incredibly influential in shaping me into the linguist I am today. His passion for this research is exceptional, and I will consider myself blessed if I can maintain half the excitement he has for this pursuit of knowledge.

Special thanks goes to the members of my committee—Robert DeKeyser, Naomi Feldman, Valentine Hacquard, and Colin Phillips. Their insight and encouragements have been invaluable. The unfailing optimism they have shown for this work has been infectious, reinvigorating my spirit more than I realized I needed.

I have received an amazing education in education, and for that I have to thank Peggy Antonisse, Tonia Bleam, and Howard Lasnik. These people pour their heart into their teaching, and it is evident in the incredible caliber of not only their classes, but also their students. They have engendered in me a love for teaching I never expected but will always value. I am especially grateful to Peggy Antonisse, who has been not only an amazing mentor but also a true friend. Her passion for leading students through to that “aha moment” and her striving to continually improve is a constant inspiration.

The University of Maryland would have never seen my face without the strong push from my undergraduate mentors, Alan Munn and Cristina Schmitt. It is their constant enthusiasm and excitement that kindled my own excitement for this research. I am grateful for them seeing my potential before even I could.

I am heavily indebted to the research staff at the Project on Children’s Language Learning. The research presented in this dissertation would be largely incomplete without the tireless efforts of many undergraduate research assistants. Special thanks are due to Mike Feters, Emma Nguyen, and Morgan Moyer, who have all given countless hours to this project. Our lab manager Tara Mease is the grease that moves the wheels. To all of you, thank you.

I am exceedingly lucky to have had the privilege of sharing this experience with the seven other students in my cohort: Kenshi Funakoshi, Angela He, Yakov Kronrod, Sol Lago, Darryl McAdams, Dan Parker, and Alexis Wellwood. I learned as much from each of you as I did in any class.

My family, especially my parents, has been an endless fountain of support, even though (or maybe especially because) graduate school put far more miles between us than they would like. With every step I am more aware of the long-attested truth that you become more and more like your parents as time passes. I am therefore incredibly lucky to be blessed with fantastic parents. I have never been more grateful to inherit my father’s strong-willed determination or my mother’s forthright sincerity. It is their strength that guides me forward.

Throughout the last five years I have grown my own family here in Maryland. I have gained the life-long friendship of so many wonderful individuals, and I am incredibly blessed to be inspired by them every day. Kate Harrigan (along with her husband, Jeff Vitols) has been a constant source of encouragement, and without her support this dissertation may have never been completed. I’ve had the amazing opportunity to work alongside her as a researcher, teacher, event planner, cook, bargain shopper, and in so many other capacities that I can easily say she has enriched every aspect of my life. Mike Feters has made me cry with laughter more times than I can count, which is a gift I cherish more than most anything else. I owe him an incredible debt of gratitude for his friendship and support (and for staying up until ungodly hours of the morning to watch me defend this dissertation from halfway around the globe). Alix Kowalski has endlessly indulged my Michigander roots. Over the hundreds of miles

between D.C. and Detroit as my road trip partner, and the tens of miles around the track as my running partner, I gained a friendship that I will treasure forever.

Although I can't elaborate on the contribution of each individual to this incredible experience, I would be remiss if I did not also thank Dustin Chacón, Michaël Gagnon, Chris and Ivana LaTerza, Darryl McAdams, Morgan Moyer, and Aaron White. Your friendships have given me an embarrassment of riches. You have each made my time here profoundly amazing and completely unforgettable. I am better for having known you.

Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
List of Tables.....	vii
List of Figures.....	viii
1 Introduction.....	1
1.1 Grammatical Restrictions on Reference Relations.....	6
1.2 Principle C in Acquisition.....	8
1.3 Goals of this Dissertation.....	13
1.4 Outline of the Dissertation.....	17
2 Heuristics for Interpreting Transitive Syntax are Not Sufficient.....	21
2.1 Outline.....	21
2.2 Introduction.....	22
2.3 Experiment 1.....	41
2.3.1 Method.....	41
2.3.2 Results.....	47
2.3.3 Discussion.....	48
2.4 Experiment 2.....	51
2.4.1 Method.....	52
2.4.2 Results.....	53
2.4.3 Discussion.....	54
2.5 General Discussion.....	55
2.6 Conclusion.....	56
3 A Bias Against Backward Anaphora is Not Sufficient.....	58
3.1 Outline.....	58
3.2 Introduction.....	59
3.3 Experiment 3.....	73
3.3.1 Method.....	73
3.3.2 Results.....	79
3.3.3 Discussion.....	80
3.4 General Discussion.....	81
3.5 Conclusion.....	83
4 Identifying Knowledge Through the Contribution of Processing.....	85
4.1 Outline.....	85
4.2 Introduction.....	85
4.3 Experiment 4.....	97
4.3.1 Method.....	97
4.3.1.1 Participants.....	97

4.3.1.2 Materials & Design.....	97
4.3.1.3 Procedure.....	105
4.3.1.4 Analyses & Predictions.....	105
4.3.2 Results & Discussion.....	112
4.4 General Discussion.....	134
4.5 Conclusion.....	138
5 Conclusion.....	140
5.1 Key Findings.....	141
5.2 Conclusions.....	165
Appendix A.....	168
Appendix B.....	171
Appendix C.....	175
Appendix D.....	178
References.....	179

1 Introduction

Learning a native language is a challenge that dominates the first few years of a child's life. And yet, with little to no explicit instruction, children are able to achieve adult-like proficiency with seeming ease. Somehow, working only with subconscious, internal cognitive systems and partial, incomplete evidence about their language from the environment around them, children form generalizations about the workings of their language that largely match the adult grammatical system by the time they are about 3. The ease with which this acquisition takes place is particularly noticeable when we compare the challenge grammarians face in finding the most accurate way to describe this internal grammatical system; many researchers have spent large portions of their career identifying, specifying, and revising generalizations to account for some set of linguistic facts. In language acquisition research, we face the challenge of explaining children's seemingly effortless acquisition of a system when we know (often firsthand) that finding the right generalization to account for the facts can be surprisingly difficult.

One primary challenge to identifying children's acquisition is the asymmetry between grammatical competence and performance. As acquisitionists, our goal is to diagnose linguistic competence, that is, underlying knowledge of a particular linguistic phenomenon. However, we cannot directly observe or measure a child's competence. What we have access to is children's performance, their behavioral response to a given linguistic stimulus. We face a challenge here in that performance is not a direct reflection of competence, because it is influenced by additional factors. If we observe children's

response to a linguistic stimulus to be non-adultlike, it could be due to children lacking the necessary grammatical knowledge to drive adult-like performance. Alternatively, children could in fact have the underlying competence, but some aspect of the deployment processes required to implement this knowledge is lacking. Since we are only able to see the end result of performance, we can't straightforwardly identify what's causing the non-adult performance. Somewhat less often discussed is the same problem on the opposite end of the scale-- where we observe children's performance to be adult-like. Here, it could be that children have the underlying competence, but it could also stem from a non-adult heuristic strategy which produces the same performance as adult knowledge. The challenge we face is to find ways to accurately diagnose syntactic competence while it is mediated by performance.

In identifying the acquisition of syntactic competence, it is important not only to diagnose children's underlying knowledge, but also how this knowledge changes over time. For a given linguistic phenomenon, we can conceive of the competence for that phenomenon developing in two ways. The first type of development can be thought of as growth of knowledge- that is, children go through stages where they have incomplete or imperfect knowledge, which approximates the adult system to some extent but is not fully parallel the adult grammar. In this case, acquisition of knowledge occurs via representational change; each stage of imperfect knowledge is like stepping stones that get closer and closer to the adult system. The second type of development can be thought of as a process of recognition rather than change of the representational system. In this case, children have some underlying understanding of the representational vocabulary of the grammatical system, and development consists of identification of the mapping

between strings in the child's language and representations. Once the child has enough information about the structure of the language they are learning, they will be able to apply their pre-existing knowledge.

Identifying the type of developmental pattern for a given linguistic phenomenon is critical to our understanding of the nature of the linguistic system. However, the imperfect relation between competence and performance once again adds a challenge: since we cannot directly measure a child's knowledge state, inferences about the pattern of development are drawn from children's behavior. This indirect inference can often mask the pattern of development.

Throughout this investigation, I focus on one test case for emerging syntactic competence in young children: Principle C. As discussed in the following section in more detail, Principle C is a syntactic restriction on the set of possible interpretations of R-expressions. The acquisition of Principle C knowledge provides an ideal test case for exploring development of competence. As a constraint, Principle C is a generalization about what a sentence *cannot* mean. As such, children receive no positive evidence in the data that would help them arrive at this generalization. This illustrates the classic Poverty of the Stimulus problem identified by acquisitionists.

The problem in fact is twofold. First, there are countless generalizations one might make to account for the facts that Principle C does; many will account for a large subset of these facts, but only Principle C accounts for all of them. This challenge was illustrated by grammarians as they sought to describe the linguistic system. Various generalizations were put forth to account for the restrictions on interpretation of reference in some sentences. Each increasingly complex generalization was designed to account for

as many of the available facts as possible. As counter-examples arose, modified generalizations were put forward to account for them. But it was not until the correct generalization for Principle C was identified that a generalization could be shown to account for all of the available facts. Thus it is clear that finding the one right generalization among many can be challenging. The second problem faced in Poverty of the Stimulus cases is that the input available to child language learners is often largely consistent with many of these wrong generalizations.

As a classic Poverty of the Stimulus case, Principle C is in fact an ideal test case for examining the development of syntactic competence. To identify the developmental pattern as knowledge recognition or knowledge growth, I examine whether children utilize any alternative generalizations for interpretation before attuning to the correct one. This exploration of alternative interpretive strategies which account for some (but not all) Principle C effects also yields one strategy for inferring competence from performance. By comparing across multiple linguistic contexts, we can identify which underlying knowledge state can account for all observed performance. Results show that across linguistic contexts, knowledge of Principle C is the only interpretive strategy which correctly predicts all behavior. This line of research provides a key insight to the developmental pattern of Principle C. Results suggest that children do not ‘try out’ alternative generalizations for interpretation, but rather seem to be adult-like at the earliest studied age point. These findings fit more with a developmental pattern where knowledge is recognized rather than knowledge growth. These findings have implications for debates over the nature of Principle C in terms of both whether it is learned or

innately specified as part of UG, as well as whether it could be derived from pragmatic principles.

Another critical contribution of this work is a novel analytic approach which utilizes measures of the processes required to implement grammatical knowledge as predictors of individual variation in performance. Recall that one of the primary challenges we face is inferring competence based solely on performance, when multiple knowledge states are compatible with the same pattern of performance. In this research, I create measures of several different processes required to deploy different types of grammatical knowledge. Different types of knowledge will require different processes for that knowledge to be implemented in interpretation. By identifying different deployment processes and determining which contribute to changes in performance, we are able to narrow in on the onset of Principle C knowledge in development. This is some of the first work which allows us to target and separate out the effects of performance factors, and actually use measures of performance as a way to recognize the state of the underlying knowledge. Rather than ignoring or attempting to minimize the performance effects in a task, we employ variation in performance as a diagnostic to identify the corresponding variation in deployment of grammatical knowledge. This, in turn, allows us to make inferences about the type of grammatical knowledge being deployed.

In the remainder of this Introduction, I specify our description of Principle C and the relevant terminology, review previous findings of Principle C effects in acquisition, identify the goals of this line of research, and provide an outline of the rest of the dissertation.

1.1 Grammatical Restrictions on Reference Relations

Principle C is part of Binding Theory, a set of three structure-dependent constraints on the interpretive relations between nominal elements in a sentence. These constraints are defined in (1) as specific restrictions on where anaphora between two NPs can and cannot occur. Necessary terminology of *binding* and *c-command* are defined in (2) and (3), respectively. Principle A governs the use of anaphors (reflexive pronouns and reciprocals), Principle B governs the use of pronominals (non-reflexive pronouns), and Principle C (1c) governs the use of R-expressions, i.e. those NPs that are not subject to Principles A or B (e.g. *the cat*, *cookies*, *Jennifer*, *every student*). The constraints of Binding Theory have been proposed to account for a wide range of facts about the types of sentences where coreference does and does not appear to be possible (e.g. Langacker, 1966; Ross, 1967; Lakoff, 1968; Lasnik, 1976; Chomsky, 1981).

- (1) a. Principle A: an anaphor must be bound in its governing category.
b. Principle B: a pronoun must be free in its governing category.
c. Principle C: an R-expression must be free.
- (2) A node α binds a node β iff:
 - a. α and β are co-indexed,
 - b. α c-commands β .
- (3) A node α c-commands a node β iff
 - a. neither node dominates the other
 - b. the first branching node dominating α dominates β .

As stated in (1c), Principle C restricts the set of possible interpretations available for sentences containing an R-expression. Specifically, it blocks a coreferential interpretation when an R-expression occurs within the c-command domain of another NP¹. In (4-6), the NP *Katie* does not occur within the c-command domain of the pronoun *she*.

- (4) While Katie₁ was in the kitchen, she_{1/2} baked cookies.
- (5) While she_{1/2} was in the kitchen, Katie₂ baked cookies.
- (6) Katie₁ baked cookies while she_{1/2} was in the kitchen.
- (7) She_{1/*2} baked cookies while Katie₂ was in the kitchen.

Consequently, co-indexation between these two NPs does not yield a binding relation and so Principle C is satisfied. In (7), however, the NP *Katie* does occur within the c-command domain of the pronoun *she*, and so if these NPs are coindexed, Principle C is violated. Thus, *she* and *Katie* must be interpreted as disjoint in reference.

The specification of c-command as the correct syntactic relation for defining binding is discussed at length in Reinhart 1976 (see also Fiengo 1977, Chomsky 1981, Chomsky 1986). Under this analysis, interpretation of possible coreference relations between NPs inherently requires a representation of the hierarchical structure of sentences. The implication for acquisition, then, is that children must have an adequate understanding of the phrase structure of their native language in order to be able to

¹ My examples will primarily deal with pronouns in a potentially c-commanding position, but note that the restriction Principle C places on coreference extends to other NPs as well.

interpret relations between NPs in an adult-like manner. (As noted above, it can be challenging to distinguish this behavior driven by adult-like grammatical knowledge from behavior driven by a non-adult mechanism used as an initial interpretive strategy before adult grammatical knowledge is achieved. We explore this challenge in detail below.)

1.2 Principle C in Acquisition

Acquisition research provides critical insight into the nature of Principle C, and as such has received considerable attention, for a number of reasons. First, the constraint is stable cross-linguistically; every language displays its effects, though in some languages these may be masked by independent features of the language (Baker 1991, 2001; Phillips 2004). Further, work with 3-5 year olds on Principle C has shown children to have fairly early and robust knowledge of the constraint (for a review, see Lust, Eisele & Mazuka 1992). Crain & McKee (1985) presented children with a short story, acted out with toys, followed by sentences like (8-10). In each case, the coreferential interpretation of the sentence was true given the story context and the disjoint interpretation was false. Children were asked to say whether the sentence was an accurate description of the story. Crain & McKee report that 3-year-old children accept sentences like (8-9), showing a coreferential interpretation, only 12% of the time, while they accept sentences like (10) 73% of the time. This pattern of restricted pronominal interpretations in Principle C contexts has been widely replicated in 3-5 year olds across a number of languages (Grimshaw & Rosen, 1990; Eisele & Lust, 1996; Guasti & Chierchia, 1999/2000; Kazanina Phillips, 2000; Leddon & Lidz, 2006), with acceptance rates of coreference between 17-37.5%.

- (8) He washed Luke Skywalker.
- (9) He ate the hamburger when the Smurf was in the fence.
- (10) When she was outside playing, Strawberry Shortcake ate an ice cream cone.

While the findings that children obey Principle C by 3-5 years of age is important, such work leaves open the question of how learners come to know this constraint, as well as what the developmental trajectory (if any) of this knowledge is. Recent research has shown that children as young as 30 months are able to interpret sentences in one Principle C context in an adult-like manner (Lukyanenko, Conroy & Lidz, 2014). This research does not conclusively demonstrate knowledge of Principle C by 30 months, as there are numerous other interpretive strategies that could yield the same result (we will discuss these strategies at length in the next section). However, it does show that by 30 months interpretation in a Principle C context is restricted to the disjoint interpretation, which is at least consistent with adult-like interpretation via Principle C.

Research by Lukyanenko, Conroy & Lidz (2014) sought to determine the earliest point in development at which children's interpretations are consistent with Principle C restrictions, in order to further investigate how Principle C emerges in development. LCL targeted performance at 28-32 months. At such young ages, children generally do not have the metalinguistic abilities to make an explicit judgment about the appropriateness of a sentence as a description of a complex chain of events, as is required in Truth Value Judgment Tasks commonly used to assess interpretations in Principle C contexts in preschool children. In lieu of an explicit-choice task, LCL utilized an Intermodal

Preferential Looking Paradigm (IPLP)², in which children are presented with two images on a television screen accompanied by an utterance of a test sentence. This paradigm is based on the finding that children attend more to an image that matches their interpretation of audio that they hear than to one that does not (Spelke, 1979; Hirsh-Pasek & Golinkoff, 1996).

The IPLP provides an ideal method for probing complex linguistic knowledge in very young children in several ways. First, it requires no conscious action from the participants, as pointing or act-out tasks do; such requirements can be cognitively demanding and have the potential to obscure underlying understanding (Hamburger & Crain 1982). Additionally, this paradigm requires fewer memory and recall capabilities than other paradigms such as truth value judgment tasks, where children must remember what happened during a story in order to accurately respond during test. Finally, unlike picture matching tasks, the IPLP allows for the use of video rather than static images as a representation of dynamic events; the use of dynamic events may be ideal for research in syntactic comprehension because sentences describe events (Waxman, Lidz, Braun & Lavin 2009). In these ways, the IPLP offers a method that may more accurately capture syntactic knowledge in young children than many other paradigms; further, its decreased demands on the child allow for use with much younger children, making it ideal for testing early syntactic knowledge.

LCL presented children with pairs of images depicting self-directed events (e.g. Figure 1, left image, girl A patting herself) and other-directed events (e.g. Figure 1, right image, girl B patting girl A). Children were presented with a sentence whose direct object

² This paradigm has also been referred to as the “looking while listening” procedure by Fernald and colleagues (Fernald et al. 2008).

was either a name, as in (11), or a reflexive pronoun, as in (12). Each sentence was therefore a good description of only one of the two events (given the adult grammar).

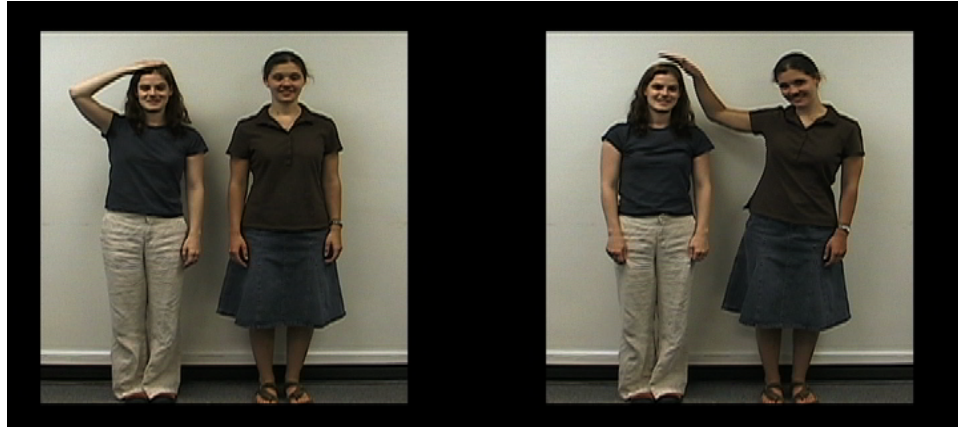


Figure 1 Principle C Task sample array from Lukyanenko, Conroy & Lidz (2014)

(11) She's patting Katie!

(12) She's patting herself!

LCL found that upon hearing (11), children looked significantly more to the other-directed action, suggesting that by 30 months children have some constraint on possible interpretations of pronouns. Further, this result was shown to be mediated by children's vocabulary size: children with larger vocabularies looked more to the other-directed event in Principle C contexts than those with smaller vocabularies. While no overall effect was found for reflexive sentences, a similar vocabulary effect was found, showing that children with larger vocabularies look more to a self-directed event upon hearing (12).

As we interpret these results, it is important to note a concern that some researchers have raised about the strength of the conclusion that can be inferred from implicit tasks such as the IPLP, which rely on measures of attention. In more explicit tasks, such as TVJTs, children's response is taken to be an explicit reflection of their interpretation of the target sentence. Comparatively, tasks based in attentional measures are not able to give evidence about which interpretations *aren't* available, only those that *are*. When children hear (11) and look more to an other-directed event, we cannot determine whether children's grammar only generates one possible interpretation, or generates multiple interpretations that children preferentially select between. The predicted looking behavior for each of these possible states of the grammar is identical. Looking more to an other-directed event in the context of a sentence constrained by Principle C indicates that children *prefer* to interpret such a sentence as disjoint rather than coreferential (for a critique of methodologies which rely attentional measures, see Crain & Thornton, 1998:55). However, it is important to note that even for explicit tasks like TVJTs, the strongest licensed claim that can be made is that children preferentially access one interpretation of the test sentence and respond to the test sentence based on that interpretation; it is virtually impossible to assure that a given interpretation is completely unavailable, irrespective of the methodology used (Musolino & Lidz, 2006; Conroy et al., 2009; Syrett & Lidz, 2011).

The preference for a disjoint interpretation in Principle C contexts shown by LCL mirrors effects shown in older children and adults. This research is some of the first evidence that children younger than 3 years old reliably show the same restricted set of interpretations in binding environments as adults do.

1.3 Goals of this Dissertation.

Given the findings on children's early interpretations in Principle C contexts, my primary goal in this dissertation is to characterize more explicitly the onset of knowledge of Principle C in children's grammar. To accomplish this, I address several questions which explore children's early behavior as a means to identify the nature of the underlying knowledge.

Question 1: At what point in development are Principle C effects observable in children's behavior?

My initial investigation will explore children's performance in Principle C contexts at increasingly early ages so as to determine the earliest stage at which we observe a restriction on interpretation. In order to characterize the developmental trajectory of Principle C performance, we need to identify the onset of such behavior. Lukyanenko, Conroy & Lidz (2014) have extended the age of onset down to 30 months, showing that Principle C behavior is in place at this point. I will investigate children at even earlier ages to determine the contrast point when children do not evidence a restriction on interpretation in Principle C contexts.

Question 2: Are early Principle C effects attributable to knowledge of Principle C, or does this early behavior arise from an alternative interpretive mechanism?

While LCL showed that at 30 months children's interpretations are consistent with knowledge of Principle C, this is not equivalent to showing that children have knowledge of Principle C at 30 months. While these results are consistent with behavior

driven by Principle C, they are also consistent with behavior driven by a number of non-adult constraints on interpretation. From the current data, it is not possible to distinguish the exact nature of the underlying constraint motivating children's preferences. Listed in (13) are several possible constraints on pronoun interpretation, all of which could yield identical behavior in this initial task.

- (13) A pronoun may not co-refer with...
- a. any NP in the same sentence.
 - b. any NP that it precedes.
 - c. any NP that it precedes in the minimal clause containing it.
 - d. any NP in the minimal clause containing it.
 - e. any NP within its c-command domain.

In the minimal sentence structure utilized by LCL (e.g. *She's patting Katie*), all of the options listed in (13) are equally satisfied. LCL's findings definitely indicate that children have some constraint on pronoun interpretation by 30 months. However, the data are consistent with myriad interpretive constraints, only one of which corresponds to adult-like Principle C knowledge. Alternatively, children could have no constraint on pronoun interpretation, but independent features of how children understand sentences could create preferences for particular interpretations.

To narrow in on the nature of the knowledge that drives children's early interpretations in Principle C contexts, I utilize two approaches. The first is to identify specific alternative interpretive strategies that have been proposed to account for early

Principle C effects and explore performance in linguistic contexts where knowledge of Principle C and knowledge based in this interpretive strategy predict differing performance. The second approach is to take advantage of the inherent dependency between grammatical knowledge and the deployment processes required to implement this knowledge; recognition of this dependency allows us to generate predictions about variance in behavior based on variance in implementation. We can draw conclusions about children's grammar by acknowledging that differing knowledge states will require different mechanisms for implementing that knowledge, and by finding ways of explicitly measuring the online implementation of such knowledge.

Question 3: What are the processing mechanisms through which children's knowledge is deployed?

As identified above, understanding what factors contribute to the deployment of children's knowledge will allow us to infer the nature of that knowledge. I explore multiple measures of children's speed of processing information in order to identify which type(s) of processing contribute to explaining performance. From this relation between deployment processes and performance, I then make inferences about what the underlying knowledge state might be.

Question 4: How did children arrive at the knowledge state that drives their early Principle C effects? Can we contribute to the innate vs. learned debate?

Although behavior consistent with adult-like interpretation is evident by 30 months of age, a constraint on possible interpretations such as Principle C poses an

interesting learnability problem. Positive evidence for a constraint is not possible- there will be no sentences produced by adults that explicitly exhibit the fact that a sentence *cannot* have a particular meaning. Additionally, evidence for a constraint on interpretations would require recognition not only of the particular sentence forms that are implicated, but also the meaning attached to a particular form. Because sentences over which Principle C applies are acceptable with disjoint reference, the necessary evidence that Principle C exists would have to come as indirect negative evidence that coreferential interpretations are blocked (Lidz, 2007). In general, the evidence for Principle C's universality across languages and robust effects in young children (along with the assumption that the input is not sufficient for the acquisition of such a constraint) have been considered evidence that Principle C is innately specified as part of Universal Grammar (UG) (Crain, 1991). Crain (1991) identifies these aspects of Principle C as "hallmarks of innateness" and has argued that by its very nature as a constraint, Principle C must be innate.

While the research presented here does not directly contribute to the debate surrounding the nature of such constraints, it does provide evidence about two aspects of the acquisition pattern which could potentially further restrict a learning account of Principle C knowledge. First, this work continues to extend downward the age at which Principle C effects have successfully been demonstrated to even younger children than the majority of previous research. This narrowing of the age range when Principle C may become active in the child grammar effectively serves to limit the amount of time across development in which the constraint could be learned. Second, this research explores the mechanisms by which children at these youngest ages implement the knowledge

responsible for their interpretations, with the aim to diagnose early Principle C behavior as stemming from knowledge of Principle C rather than an alternate interpretive heuristic. In this way, demonstration of children's early accurate interpretations in Principle C contexts could serve to constrain the amount and type of data available to drive learning, and hence place some empirical bounds on arguments for or against its innateness.

Question 5: Do children's early Principle C effects contribute to identifying the nature of Principle C as a primitive of grammar?

While Principle C is commonly identified as a primitive of grammar, alternative approaches to account for Principle C effects have proposed that the restriction on interpretations derive from discourse pragmatic principles (Reinhart, 1983; Grodzinsky & Reinhart, 1993; Ambridge, Pine & Lieven, in press). The results of our explorations of children's early interpretation in Principle C contexts are relevant to this debate in that whatever pragmatic competence is required for such interpretive processes to be implemented must necessarily be available to children by 30 months. We discuss the likelihood of this possibility given current findings on children's early abilities in the relevant pragmatic domains.

1.4 Outline of the Dissertation

Chapter 2 examines the possibility that children's early interpretations in Principle C contexts could be attributable to heuristics for interpreting transitive structures rather than to knowledge of Principle C. Research in the word-learning domain suggests that children utilize cues from argument structure to identify a novel verb's possible

interpretations. This interpretive strategy from word learning provides an alternative mechanism to Principle C by which children could develop a preference for disjoint interpretations in simple Principle C contexts. In two preferential looking tasks, I explore performance by 24 and 30 month-olds to diagnose the interpretive strategy responsible for early Principle C effects. I utilize a novel linguistic context where knowledge of Principle C and a transitivity bias predict differing performance.

The results contribute two significant findings. First, I confirm that the youngest age at which Principle C effects are observable in children's behavior is 30 months; no significant restrictions on interpretation are evident at 24 months. Second, I show that in a new context where Principle C and a bias based in transitivity predict differing performance, behavior is more consistent with Principle C. This suggests that early Principle C effects are not attributable to an alternative interpretive strategy based in transitivity.

Chapter 3 explores the possibility that children's early interpretations in Principle C contexts could be attributable to a mechanism that relies on the linear order in which a pronoun and an R-expression occur. One possibility is that children could begin with a grammatical rule which disallows all cases of backward anaphora as an initial attempt at interpreting reference relations. There are many possible generalizations about licensed interpretations Principle C contexts which account for most of the facts (although, as noted above, only one that accounts for all of the facts). A bias based in the linear order of nominal expressions is one example of a wrong generalization which can account for interpretations of most of the simple sentences children hear. As such, it provides an

excellent test case to help determine whether children test out alternative generalizations before settling on the correct generalization (i.e. Principle C).

In Chapter 3 I compare 30 month-olds' interpretations of two types of sentences where a linear bias forms different predictions than does Principle C. I utilize backward anaphoric sentences where a pronoun c-commands its potential antecedent and a coreferential interpretation is blocked by Principle C, and backward anaphoric sentences where this c-command relation does not obtain, and thus a coreferential interpretation is grammatically available. Children's performance is consistent with knowledge of Principle C, and not a linear bias, in these more complex frames. This shows that even at increasingly young ages, children do not utilize alternative interpretive mechanisms.

Chapter 4 examines a novel analytic approach to identifying characteristics of knowledge underlying behavior. I utilize independent measurements of processes that may be implicated in deployment of knowledge in order to determine the nature of the knowledge that is being deployed. I exploit the predicted dependency between children's speed of processing syntactic information and their speed of interpreting sentences in Principle C contexts. In three preferential looking tasks, I use measures of processing at the lexical and syntactic levels to compare to performance on a Principle C task. I show that individual variation in the speed of interpretation in Principle C contexts is predicted by individual variation in syntactic processing speed. This finding suggests that the mechanism responsible for interpretation in Principle C contexts is dependent on syntactic composition.

Chapter 5 summarizes the findings identified in Chapters 2-4 and explains how these findings answer the questions posed in this Introduction. I discuss how the research

presented in this dissertation has strengthened our understanding of the developmental pattern of Principle C in acquisition. I examine how these findings serve to constrain any proposed learning account of Principle C, and provide an initial investigation of children's language input to evaluate the plausibility of learning from such input. I also explore several interpretive mechanisms based in discourse pragmatics which have been proposed as alternatives to Principle C, and I examine the efficacy of these interpretive mechanisms in light of the findings presented herein.

2 Heuristics for Interpreting Transitive Syntax are Not Sufficient

2.1 Outline

Chapter 2 examines the possibility that children's early interpretations in Principle C contexts could be attributable to heuristics for interpreting transitive structures rather than to knowledge of Principle C. Research in the word-learning domain suggests that children utilize cues from argument structure to identify a novel verb's possible interpretations. This interpretive strategy from word learning provides an alternative mechanism to Principle C by which children could develop a preference for disjoint interpretations in simple Principle C contexts.

This chapter addresses two primary questions. First, what is the youngest point in development when children exhibit Principle C effects? In order to develop a complete understanding of the learning profile of this constraint, it is imperative to establish when performance is consistent with the constraint. Second, are these apparent Principle C effects driven by knowledge of Principle C at this earliest point in development, or might another mechanism be responsible for this behavior? Findings suggest that the onset of Principle C behavior appears sometime between 24 and 30 months, and is in place by 30 months.

2.2 Introduction

Children's treatment of simple sentences such as (14) have been shown to mirror adults' in disallowing a coreferential interpretation as young as 30 months. (Lukyanenko, Conroy & Lidz, 2014).

(14) She's patting Katie.

Lukyanenko, Conroy & Lidz (2014) interpret this behavior as suggestive that children have a syntactic constraint on interpretation such as Principle C, which prohibits coreference between an R-expression and a c-commanding pronoun. However, it is important to remember that observed behavior is not a direct reflection of the underlying knowledge driving that behavior. While this behavior pattern is consistent with children having adult-like knowledge of Principle C, it is equally consistent with numerous other potential strategies for interpretation. Recall that there are myriad alternative generalizations one might make which can account for *most* of the interpretive facts which Principle C correctly captures. These alternative generalizations only approximate the interpretive mechanism available in the adult grammar, not being able to account for all of the data. However given the simple single clause structures used by LCL, they are just as accurate at driving a disjoint interpretation as Principle C. Thus even in the absence of understanding the correct restrictions on reference relations which are employed by the adult grammar, children would have access to a wealth of mechanisms that could drive their early interpretations.

In this chapter I examine children's early interpretations of similar sentences as observed in word learning tasks, and identify a pattern that has been attested in this literature which suggests a possible alternative mechanism by which children could formulate a preference for a disjoint interpretation to sentences like (14). Studies of syntactic bootstrapping in young children have shown that children routinely exhibit two critical expectations about the type of event that corresponds to a sentence given its structure. First, children expect the number of arguments in a clause to be matched to the number of participants in the event. Second, children expect that a sentence with a transitive frame will correspond to an event with a causal interpretation. Together, these two interpretive biases exhibited by young word learners could yield a disjoint interpretation in the case of sentences like (14) just as easily as adult grammatical knowledge of binding constraints. In the simple context of sentences with a *X VERBs Y* form, both of these interpretive mechanisms are equally consistent with the behavior observed in 30 month-olds.

This chapter explores the nature of the underlying knowledge that yields disjoint interpretations of sentences like (14) at 30 months. In order to determine which mechanism is responsible for interpretation, I compare performance predicted by adult-like understanding of Principle C and an alternative strategy based in biases for transitive syntax. Given the findings identified above, I specify two possible variants of a bias that relates transitive structures to two-participant events. While these mechanisms predict identical performance for simple sentences like (14), I utilize a linguistic context in which each mechanism predicts a different performance pattern in order to identify which is responsible for children's behavior. I take advantage of a critical difference between

these interpretive mechanisms- the relevance of the type of nominal element of each argument. Interpretation via Binding Theory varies given the type of nominal elements in each argument. Unlike Binding Theory, however, a Transitivity Bias does not differentiate between nominal types, depending only on number of arguments. I therefore compare children's interpretations for sentences like (15) and (16).

(15) She's patting Katie's head.

(16) She's patting her head.

Because they do not differentiate between NPs with a phrasal possessor and NPs with a pronominal possessor, both versions of a bias in transitive clauses predict interpretations to be identical across both sentence types. However, because Principle C is formulated with respect to the distribution of R-expressions, it applies only to (15) and therefore only predicts a restriction on interpretation in one case.

In the current study I compare children's responses to sentences like (15) and (16). An asymmetry in interpretations of these two sentence types is expected by Principle C but not by either formulation of a transitivity bias. I show that children's response patterns differ to each of these sentences, suggesting that an interpretive strategy based in transitivity could not account for the performance exhibited here and is thus not likely to be the knowledge driving behavior in these contexts.

For the remainder of this chapter, I will begin by identifying relevant observations from previous research on syntactic bootstrapping and describe two possible formulations of an interpretive strategy based on the biases identified in the literature. I will then

present an overview of the current study which was designed to address the contribution of the knowledge driving performance to simple sentences in Principle C contexts.

Interpretive strategies for transitive structures

One mechanism that has been shown to drive children's interpretations of sentences similar in form to (14) has been suggested in the word learning literature. Given that children must form an interpretation of a sentence with a verb they have not heard before in order to learn the meaning of the verb, researchers have suggested that word learners use the argument structure to make inferences about the possible meaning; this process is called *syntactic bootstrapping* (Landau & Gleitman, 1985; Gleitman, 1990). Syntactic bootstrapping inferences rely on the fact that there is a systematic relation between the types of syntactic structures a verb may be used in and the types of meanings a verb may have (Fisher, Gleitman & Gleitman, 1991; Jackendoff, 1983, 1990; Levin, 1993; Naigles, Gleitman & Gleitman, 1993, inter alia). Categories of verbs can be described by the number and type of participants a verb specifies. For example, verbs that describe one individual's action on another are likely to appear in a transitive syntactic frame with two NP arguments, and unlikely to appear with a sentential complement, as this is not a role licensed by the verb category. We may recognize numerous systematic generalizations about the set of frames that a verb can occur in and the meaning of the verb. It is this systematic relation between the semantic type of a verb and the set of syntactic structures it licenses which is responsible for the intuition that the novel verb *blick* in (17) below describes some asymmetric action of one participant on another and the novel verb *frump* in (18) describes a mental state.

- (17) John blicked Bill.
- (18) Mary frumped that Bill was asleep.

Investigations into children's ability to utilize syntactic bootstrapping as a mechanism for learning novel verbs has shown that children use syntactic information to classify verbs from very early ages- beginning even before their second birthday (Arunachalam et al., 2013; Gertner, Fisher & Eisengart, 2006; Naigles, 1990). Further, children have been shown to use a number of different syntactic cues as a means of interpreting sentences with novel verbs. By 28 months, children predict a match between the numbers of syntactic arguments licensed by a verb and participants involved in a corresponding event. They interpret novel verbs presented in transitive frames as referring to two-participant events and verbs presented in intransitive frames as referring to one-participant events (Yuan & Fisher, 2009). 21 month-olds have been shown to link the agent role of an event to the subject position of a sentence, reliably interpreting sentences with transitive frames as referring to events where the depicted agent matched the NP subject (Gertner, Fisher & Eisengart, 2006). As children develop, they are able to use the animacy of a subject as a cue to the causative nature of a novel verb. Both causal alternation verbs like *break* and unspecified object verbs like *clean* can appear in the same syntactic frames, as in (19-20); however, the verb types differ in which semantic role is assigned to the subject in intransitive frames.

- (19) a. Anna broke the lamp.
b. It broke.
- (20) a. Anna cleaned the lamp.
b. Anna cleaned.

By 28 months, children familiarized with novel verbs with animate intransitive subjects interpret the novel verb as referring to a causative event while those familiarized with inanimate intransitive subjects interpret the verb as referring to a non-causative event (Bunger & Lidz, 2004, 2006; Scott & Fisher, 2009).

Most relevant to our concerns here are the findings which have shown that children form an early link between transitive syntax and two-participant events. Naigles (1990) pioneered the exploration of this link in 24 month-olds. After familiarization to a two-participant event that had both an asymmetric component (one participant acting on another- a duck pushing a bunny into a squatting position) and a simultaneous component (two participants performing the same action at the same time- the duck and bunny waving their arms), children were presented with two events simultaneously which depicted each component of the familiarization event separately. Results showed that children who heard a description of the familiarization event with a transitive frame as in (21) interpreted the novel verb as referring to the asymmetric component during test, while children who heard a description with an intransitive frame as in (22) interpreted the verb as referring to the simultaneous component.

- (21) Oh look! The duck is gorging the bunny!

(22) Oh look! The duck and the bunny are gorp-ing!

Children consistently expect a verb presented with transitive syntax to refer to a two-participant event. This pattern has been robustly attested throughout the verb-learning literature (Arunachalam & Waxman, 2010; Arunachalam et al. 2012; Hirsh-Pasek & Golinkoff, 1996; Naigles, 1990; Naigles & Kako, 1993; Noble, Rowland, & Pine, 2011; Yuan & Fisher, 2009; *inter alia*).

The bias for children to assume a causative interpretation for transitive syntax has been shown to pervade cross-linguistically, even when it is not the most reliable cue to causativity. Lidz, Gleitman & Gleitman (2003) demonstrated that Kannada-speaking children utilize the cross-linguistic cue of the number of overt arguments more readily than a language-specific causative morpheme, even though the latter is statistically a better predictor of a causative interpretation. Kannada, a Dravidian language of southwestern India, exhibits a causative morpheme that obligatorily occurs in any sentence where a causal interpretation is intended, with the exception of a limited set of verbs that allow causativity to be expressed lexically. Lidz, Gleitman & Gleitman presented Kannada-speaking children aged 3;6 with known motion verbs which occurred in transitive frames but lacked the obligatory causative morpheme (which yields such sentences to be considered ungrammatical by adult speakers). They showed that children infer a causal interpretation for these motion verbs when presented in transitive syntax.³

Thus cues such as the number of arguments in a structure seem to be one highly salient

³ Note that ‘transitive syntax’ in this context refers to the number of arguments present in the sentence, although the verbs used in this study were intransitive verbs marked with causative syntax, e.g. of the form *X Y fall-caus* to give the interpretation that X caused Y to fall.

cue to a causal interpretation to young children, even when it is to the apparent detriment to adult-like judgments of Principle C contexts.

We may conclude from these findings that from a very young age, children use cues from verb syntax as mechanisms for interpreting sentences and relating them to events. Specifically, children recognize from an early age that the number of NP arguments (usually) matches the number of event participants. This link between arguments and participants biases children to treat transitive sentences of the form *X VERBs Y* as referring to a particular type of event, namely one where an agent *X* performs an (asymmetric) action on a patient *Y*.

These early transitive syntax biases present a potential challenge to the conclusions drawn by Lukyanenko, Conroy & Lidz (2014). Because the simple sentences used in their experiment are of this same *X VERBs Y* form, it is possible that the behavioral patterns exhibited could be the result of an interpretive strategy that relies on these transitive syntax biases, rather than adult-like interpretation based in constraints on the reference relations between the R-expression and a c-commanding pronoun. That is, while knowledge of Principle C predicts a disjoint interpretation in Principle C contexts like those tested by LCL, the fact that the observed behavior is consistent with predicted behavior of this hypothesis is not sufficient to conclude that Principle C is the mechanism that drives interpretation at 30 months.

Consider the alternative: that children have no *structural* constraint on interpretation of R-expressions. In order to choose between two possible interpretations of a sentence normally constrained by Principle C, children could potentially turn to a non-structural mechanism that restricts interpretations. The simple clause structure tested

by LCL is identical to that in which the transitivity bias is most useful. When children hear the string *she's patting Katie*, they could be responding to the transitivity of the verb rather than the structure of the clause. If children applied the bias for interpreting transitive clauses here, then this string would be interpreted as corresponding to an event of one girl patting another. In the minimal visual context provided by LCL, this interpretation matches the other-directed action but not the self-directed action. So in this simple context, a bias to interpret transitive structures as referring to two participant events predicts a disjoint interpretation just as well as Principle C knowledge would. Given the simple linguistic context utilized by LCL, we are unable to differentiate between these two hypotheses, as the observed interpretations are equally consistent with either of these possible underlying knowledge states.

Beyond recognizing this alternative account for this behavioral pattern, we might also consider that children might in fact be likely to employ biases if they lack some other mechanism for choosing an interpretation, such as Principle C restrictions. Children have been shown to rely on the clause structure, even over prior knowledge of a verb's distribution, as late as 3;6-- far later than LCL's observations of 30 month-olds (Naigles, Gleitman & Gleitman, 1993; Lidz, Gleitman & Gleitman 2003, 2004). Thus treating all transitive syntax as relating to two-participant events is an inherently strong bias, and definitely a live option as an interpretive strategy for LCL's simple sentences in the absence of another mechanism which restricts interpretation.

In the present study, I address the question of whether a disjoint interpretation for sentences in Principle C contexts could be the result of an alternative mechanism for interpreting transitive syntax. In order to identify which underlying mechanism drives

this behavior, it is necessary to identify contexts in which the two mechanisms we wish to compare would predict differing performance. To specify such contexts, I will first concretize our descriptions of the alternative mechanism we have hypothesized so that the predictions such a mechanism might make are clear.

I consider two primary ways in which a bias that results in causative interpretations for transitive clauses might be formalized (although there could certainly be others). The difference between these mechanisms can be thought of as a difference in the precise nature of the expectation children exhibit. One possibility is that children form an expectation about the relation between argument structure and event structure; I will refer to this as the Event-level Transitivity Bias, defined in (25). Given this representation of the bias, children expect the number of arguments in a clause to have a 1:1 match to the number of participants in an event described by that clause. In the case of a transitive clause, this bias would cause children to predict an event depicted with two participants, one for each of the two syntactic arguments in the transitive clause. Another possibility is that children form an expectation that each NP in a clause refers to distinct entities, and this expectation manifests in their interpretation and mapping to event structure; I will refer to this as the Sentence-level Transitivity Bias, defined in (26). This specification of the bias results in the same predictions for a transitive clause; children would expect that the entity denoted by the subject and the object are distinct, and thus (assuming both participant roles corresponding to these arguments are represented in the event) predict an event with two participants.

- (25) Event-level Transitivity Bias: the number of arguments in a clause must exactly match the number of unique participants in an event described by that clause.
- (26) Sentence-level Transitivity Bias: the arguments in a clause must each refer to distinct entities.

So although each of these representations of an interpretive bias for transitive syntax⁴ manifest in the same expectation that the syntactic arguments and the event participants are in a 1:1 relation, they arise from different sources. The Event-level Transitivity Bias would arise from an assumption about how event participants relate to the syntactic arguments they are described by, and the Sentence-level Transitivity Bias would arise from an assumption about how syntactic arguments relate to each other.

Having defined our two formulizations of a transitivity bias, we can identify how they might predict performance in contexts such as those tested by Lukyanenko, Conroy & Lidz (2014). Recall that children were presented with sentences such as (27), along with two scenes that corresponded to a coreferential and disjoint interpretation of the sentence; in this case, children saw a scene where Katie was patting herself, and one where Anna was patting Katie.

- (27) She's patting Katie.

⁴ While I have been discussing these interpretive strategies in terms of their relating to their use with transitive clauses, it should be evident that in fact these strategies are equally applicable to the interpretation of intransitive clauses, and the predicted performance is consistent with the findings for intransitive structures discussed above. The focus on transitive syntax and the application therein is simply due to this being the type of structure in which Principle C effects are observed, and thus where the comparison is relevant.

Using the Event-level Transitivity Bias, children would recognize that the clause contains two arguments- *she* and *Katie*. The Event-level Transitivity Bias would prompt a search for a representation of the event which contained two corresponding participants. In this case, the scene where Anna and Katie are both participants in the patting event is a more appropriate match. The Event-level Transitivity Bias predicts a preference for the scene where Anna is patting Katie- the disjoint interpretation.

Using the Sentence-level Transitivity Bias, children would recognize that the clause contains two arguments- *she* and *Katie*. These two NPs are recognized as both referring to individuals; the Sentence-level Transitivity Bias dictates that they must refer to distinct individuals, so the event described must contain two distinct participants. Given the scenes presented, the one where Anna and Katie are both participants in the event is therefore the best match to the description. The Sentence-level Transitivity Bias predicts a preference for the scene where Anna is patting Katie- the disjoint interpretation.

Now that we have specified exactly what the mechanistic process might be for each of these interpretive strategies, it is clear that the predicted behavior for both is consistent with the behavior observed by Lukyanenko, Conroy & Lidz (2014). Given a simple transitive clause such as (27), both versions of a Transitivity Bias expect a preference for the disjoint interpretation. In this context, a generalization which stems from verb transitivity accounts for the data just as well as adult grammatical knowledge. Thus either of these alternative interpretive strategies could be the source of a disjoint interpretation for these simple sentences in Principle C contexts, rather than knowledge

of Principle C itself. An exploration of the developmental pattern of Principle C effects will allow us to determine whether children can be shown to employ alternative generalizations before hitting on Principle C.

As we discuss these alternative interpretive strategies, it is important to recognize that while these biases predict the correct interpretation a large portion of the time, they are by no means effective in every circumstance in leading children to an accurate interpretation. These alternative generalizations approximate the adult grammar, but cannot account for all the data that Principle C does. The bias to interpret transitive syntax as corresponding to an event which includes two participants is one that will lead to an accurate interpretation *most* of the time, and thus is statistically reliable enough to be an effective strategy employed in the context of narrowing the field for the identifying the meaning of a novel verb or the meaning of a clause. However, there will be a certain amount of the data for which this simplistic heuristic is not accurate.

Sentences with reflexive pronouns, like (28) present a challenge, because their meaning is determined through coreference with another argument within the clause.

(28) She's patting herself.

The coreferential interpretation intended with reflexive pronouns directly contradicts the Sentence-level Transitivity Bias, which predicts that all arguments will refer to distinct entities. Accurate interpretation of reflexive pronouns using an Event-level Transitivity Bias would require children to recognize that a single entity could satisfy multiple participant roles (for example, that one individual is both the agent and

patient of a washing event in a sentence like *the girl washed herself*). However, this exception contradicts the cue of a 1:1 mapping between arguments in a clause and participants in an event (and individuals represented in a visual portrayal of that event) which makes this bias such a strong interpretive mechanism. Under either specification of a transitivity bias, reflexive pronouns are a challenge to relying only on this interpretive strategy.

Additionally, to achieve the adult-like behavioral pattern observed by LCL would require that children recognize reflexive pronouns as the critical cue to this doubling-up of thematic roles. Recall that LCL found distinctly different response patterns to sentences like (27) and (28): children looked more to an event of another character patting Katie when hearing (27), and more to an event of Katie patting herself when hearing (28). This asymmetry in responses to two types of transitive sentences suggests that a transitivity bias alone could not account for these results. While some form of a transitivity bias could be used to arrive at the interpretation for sentences like (28), we cannot account for the behavior shown in interpreting reflexives without some additional interpretive mechanism. Children would necessarily need some understanding of reflexive pronouns and the fact that transitive sentences with reflexive pronouns differ in terms of the way in which participant roles are parceled out in the corresponding sentence.

Another challenge to interpretation via these mechanisms relates to the notion of event participants and the mapping between arguments and participants. Research by He and colleagues has identified several cases where the 1:1 match between NP arguments in a clause and understood participants in an event does not hold. If the 1:1 match for

arguments and participants does not hold universally, then it becomes a less effective interpretive strategy. Thus exploration of these cases is critical to understanding the strength of these interpretive mechanisms. In one case, He et al. (2014) showed that 10 month-old infants represent an instrument used in an opening event as a privileged event participant. However, most words that could be used to describe such an event (e.g. *the girl opened the box*) do not express the instrument as an argument. Another case in which intuitive event participants are not represented in argument structure is evidenced in resultative constructions of Mandarin, in which the patient of an event is not expressed syntactically by an argument (e.g. *John wiped the cloth dirty*, meaning that John wiped *something* with a cloth until the cloth became dirty). Mandarin children accurately represent this interpretation by 2;6 (He et al., 2014). In order to determine whether these cases of an apparent lack of the 1:1 match are problematic, the notion of event participant will need to be more carefully specified.

The present study

The goal of the current study is to identify and test a context in which the transitivity biases predict differing behavior than knowledge of Principle C, so that we may be able to diagnose which of these sources is responsible for interpretations of sentences like (27) at 30 months. Separating out behavior from the source of knowledge driving that behavior is critical to building an accurate understanding of the developmental pattern of Principle C. Should the results suggest that performance is most consistent with utilizing this transitivity bias strategy, then it suggests necessarily that adequate knowledge of Principle C is either as-of-yet unavailable to children at 30

months or that some aspect of the task precludes their ability to effectively deploy this knowledge for use in interpretation. This research seeks to identify whether children begin with an initial interpretive strategy besides knowledge of Principle C for computing reference relations. A better understanding of this developmental pattern will be one of the first steps to identifying the source of Principle C knowledge as a component of the grammar.

Now that we have specified two mechanisms by which this transitivity bias might be formed, we can narrow in on contexts in which they might predict non-adult behavior. In order to identify the knowledge which drives interpretation of simple transitive sentences in Principle C contexts at 30 months, I will examine two types of transitive sentences, identified in (29) and (30) below.

(29) She's patting Katie's head.

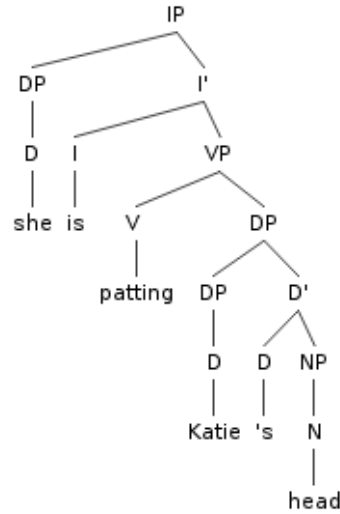
(30) She's patting her head.

Two properties of these sentences will allow us to distinguish between the interpretive strategy which yields behavior in young children. First, the transitive syntax which drives interpretation via a transitivity bias is held constant across both sentences. Because of this, either specification of the transitivity bias should predict identical performance across both sentences. However, the arguments in the clause are manipulated such that Principle C applies in the case of (29) but not (30); if Principle C is the interpretive mechanism driving performance, it predicts a divergence in performance between these sentence types. Thus recognizing whether performance varies with

sentence type should allow us to distinguish between the use of a transitivity bias or Principle C as the interpretive mechanism. (Although there are certainly other alternative mechanisms for interpreting such sentences than simply these two options, as I will discuss in Chapter 3, I will consider Principle C as the primary alternative to the transitivity bias for simplicity here.)

The second property of these two sentences, which will allow us to distinguish between predictions made by each version of the transitivity bias, is the types of entities that each argument refers to. The subject in both cases is the pronoun *she*, which can be taken to refer to an individual; alternatively, the object of both sentences refers to a component of an individual- a body part. Only one of the two formulations (the Sentence-level Transitivity Bias) takes into account the relation that the syntactic arguments bear to each other. The overall behavioral response to both of these sentences should therefore allow us to identify whether the relation of the arguments to each other is acknowledged in interpretation or not, identifying which formulation of the transitivity bias drives interpretation (if indeed either does). I will now consider each of these interpretive strategies individually to lay out the predictions that it makes for each sentence.

Beginning with Principle C, we can identify immediately that it will not apply in the case of a sentence like (30), as there is no R-expression to apply to. This is consistent with adult's intuition that the subject pronoun *she* could refer to any salient female (in the contexts relevant here, it could refer to the patient of the patting event, denoting a self-directed action, or another girl, denoting an other-directed action). However, given the structure in (31) for the string in (29), we can see that the R-expression *Katie* is contained within the c-command domain of the subject pronoun *she*.



(31)

Given that c-command obtains, Principle C restricts *she* from being co-indexed with *Katie*, resulting in the constraint against a co-referential interpretation. This corresponds to the intuition that a sentence like (29) cannot reasonably describe a self-directed action of Katie patting her own head. Therefore if Principle C is the knowledge driving children's interpretation of such sentences at 30 months, we predict a preference for a disjoint interpretation of (29), and no restriction of interpretation for (30).

The Event-level Transitivity Bias formulated in (25) above restricts interpretations of sentences to those where the number of arguments in the clause matches the number of unique participants in the described event. Given the transitive frame used in (29) and (30), which has a subject and an object argument, such a bias would predict a preference for an interpretation of a two-participant event.

Critically, we have chosen a context where the Event-level and Sentence-level formations of a transitivity bias might manifest themselves differently, due to the nature of the relation that entities denoted by the subject and object in each sentence bear to one

another. Recall that the Sentence-level Transitivity Bias restricts the interpretation of the arguments from referring to the same entity; however, in the case of (29) and (30), the pronoun subject refers to an individual and the object NP refers to a body part (someone's head). Given the nature of the subject and object arguments as inherently referring to distinct entities, for both (29) and (30) the Sentence-level Transitivity Bias can be considered to be vacuously satisfied. As such, this bias makes no restriction on interpretation.

With only two sentence types, we will be able to recognize a distinct behavioral pattern for three different interpretive strategies. A transitivity bias predicts the same behavior across conditions- the Event-level Transitivity Bias predicts a restriction on interpretation in both cases, and the Sentence-level Transitivity Bias predicts no restriction in either case. Alternatively, accurate application of Principle C predicts a restriction on interpretation only in the case where a c-commanded R-expression is present, in (29) but not (30).

Overview of the Experiments

The goal of the current study is to answer two questions. First, what is the earliest point in development at which children exhibit Principle C effects? Second, at the earliest stage where Principle C effects are exhibited, is performance more consistent with an underlying bias in interpreting transitive structures, or with accurate knowledge of Principle C? Experiment 1 examines 30 month-olds' performance in a preferential looking paradigm which presents visual arrays corresponding to events depicting a coreferential and a disjoint interpretation of sentences such as (29) and (30). I compared

attention following the onset of the object noun phrase to determine whether children exhibited a preference for the disjoint interpretation in response to either sentence type. The results demonstrate that 30 month-old children exhibit a preference for the disjoint interpretation in response to sentences like (29) but not (30); this performance is expected only if Principle C is the source of knowledge driving interpretation, and cannot be accounted for by appealing to a transitivity bias. Experiment 2 was conducted as a follow-up on 24 month-old infants with the same procedure, in order to determine whether the same pattern is exhibited at an even younger age, and if so whether the knowledge driving that behavior is similar or could instead be attributable to a transitivity bias. Results show that children exhibit no systematic preference for a disjoint interpretation in any pattern that is predicted by our interpretive strategies. Performance at 24 months is therefore likely attributable to chance.

2.3 Experiment 1

2.3.1 Method

Participants

32 30 month-olds (16 males) with a mean age of 30;11 (range 28;3 to 31;30) were included in the final sample. An additional 2 children were excluded due to experimenter error when running subjects. Participants were recruited through the University of Maryland Infant and Child Studies Database from the College Park, MD area. Children participated on a volunteer basis, and were given a small gift for participating. All participants were learning English at home, and English input constituted at least 80% of their language input. Parents filled out a MacArthur Communicative Development

Inventory: Words and Sentences questionnaire (Fenson et al., 1994). Mean productive MCDI vocabulary was 560 words (range: 295-680 of 680 possible).

Materials & Design

Visual stimuli were identical to those utilized by Lukyanenko, Conroy & Lidz (2014). All video was presented on a 51” plasma TV screen. Videos were edited together to create the sequence of events outlined in the Design section below. A full schematic of the events children see throughout the video can be found in Appendix A. Single videos were presented in the center of the screen and were sized between 18-20 inches wide and 20-24 inches tall. Videos presented together in preferential looking format were always identical in size and scale, measuring 16 inches wide and 18 inches tall, with a 7 inch span between the inside edges of the two images.

Audio stimuli was recorded in a sound-proofed room by a female native speaker of American English using intonation common in Child Directed Speech. Recordings were edited and combined with the visual stimuli.

During the Character Introduction phase, children were presented with video clips (3.5-6.5 seconds each) describing and naming individually each of two characters (Katie and Anna), who later performed all the actions during test. In addition to occurring at the beginning of the sequence, several Character Introduction clips were placed throughout the video to provide filler material between test trials as well as to further remind children of the character names. Following the Character Introduction phase, children received six Face Check sequences (6.5 seconds each), which presented these two characters on opposite sides of the screen in a preferential looking paradigm, along with a sentence

asking them to find one of the characters (e.g. *Where's Katie? Do you see Katie?*). These Face Check sequences ensured that the children were adequately mapping the names they heard in the introductory clips to the accompanying faces and could distinguish the two characters from one another⁵. Additionally, these Face Checks also served to prepare children for the test trials, where they would be required to preferentially attend to one of two images on the screen. Target character and the side of the screen on which the target character appeared were counterbalanced across Face Check trials. Order of Face Check trials was counterbalanced across subjects.

Figure 2 presents a schematic of a typical Familiarization and Test phase sequence. Events depicted in familiarization and test trials were all continuous actions, performed with both characters on the screen (even if only one character was a participant in the event). Self-directed actions consisted of scenes with one character performing an action on herself (e.g. Figure 2 leftmost image, Katie patting her own head), with the other character present but not interacting. Other-directed actions consisted of similar scenes, with one character performing the same action on the other character (e.g. Figure 2 center image, Anna patting Katie's head).

⁵ Although note that, strictly speaking, given the specific types of sentences used in the test phase and the visual context in which they were presented, knowing the characters' names was not crucial to forming an interpretation. Thus these 'face checks' primarily served to facilitate processing of subsequent sentences.

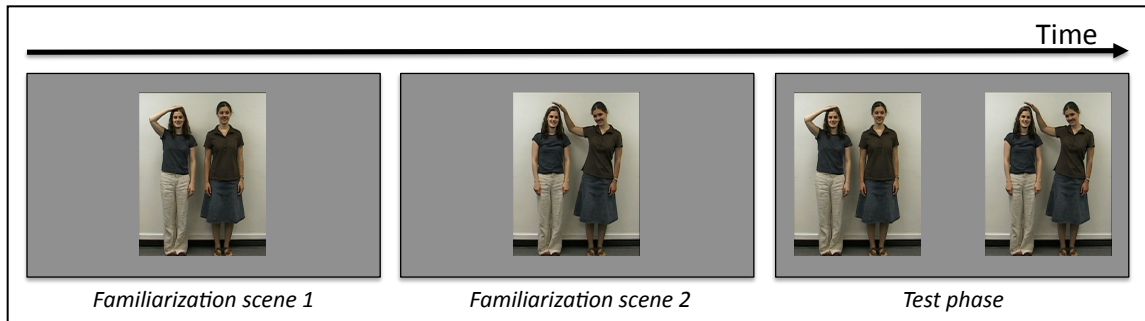


Figure 2 Schematic of Familiarization and Test phase sequence in Experiments 1 and 2.

Immediately preceding each test trial, children were presented with a Familiarization phase. In two video clips (6 seconds each), the test events were presented one at a time, with audio which described the action and the patient of the event, but which did not uniquely identify the agent of the described event, as in (32) below. This was to allow either the self-directed or the other-directed events to be a possible representation of the test sentences. Order in which the Familiarization events appeared was counterbalanced across trials.

- (32) Wow! There's Anna and Katie! It looks like Katie's head is being patted!
 Oh look! There they are again! Katie's head is being patted again!

The Test phase (12 seconds each) immediately followed with the same two event videos as shown in the Familiarization phase presented side by side simultaneously (e.g. Figure 2, rightmost image). An attention-getting filler phrase (*Oh look- now they're different!*) introduced the events. Test audio consisted of two instances of the test sentence in different frames, as shown in (33) and (34) below.

- (33) She's patting Katie's head! Find the one where she's patting Katie's head.
- (34) She's patting her head! Find the one where she's patting her head.

Children were presented with a total of 8 test trials in a between subjects design: half of the subjects heard NAME condition sentences as in (33) for all trials, and half heard all PRONOUN condition sentences as in (34)⁶. Trial order and the side of the screen each video appeared on was counterbalanced across subjects.

For each iteration of the test sentence, the audio was aligned by the onset of the object NP for later analysis, as this is the critical point when an interpretation strategy based in Principle C could begin to identify the intended interpretation of the sentence. This resulted in a 2 (Condition) X 2 (Window) design; the dependent measure was the proportion looking time attending to the other-directed action.

In addition to the phases described above, children also saw 4 filler video clips (7 seconds each) spaced throughout the video. These clips depicted children's toys presented with classical music, and were used to attract children's attention to the screen, as well as to provide a 'mental break' between the more challenging test trial sequences. For all video clips, there was 1 second black screen between each clip, and the first audio clip began 20 frames (.67 seconds) after the visual stimuli appeared on the screen.

⁶ This change from the within-subjects design used by LCL was made after researchers found children to respond more consistently in an IPLP task probing syntactic competence when they are presented with the same structure across trials than when presented with multiple sentence types (Gagliardi, Mease & Lidz, submitted).

Procedure

Participants were tested individually, sitting either in a high chair or on their parent's lap in front of the television screen. A camera mounted above the television recorded participants' eye movements during the videos, allowing more precise offline coding. Data was coded frame-by-frame using SuperCoder software (Hollich, 2005), indicating whether children were attending to the left or right side of the screen, or not at all. Coders were trained researchers who were blind to condition and could not hear the auditory stimuli. 10% of the data was coded by both coders to ensure accuracy and reliability across coders; inter-coder reliability was high: across three coders, percent agreement was above 90% in all cases, with Cohen's kappa scores of .85 and above.

Analyses & Predictions

Performance is measured over the first 3000 ms following the disambiguation point, when children's behavior is most indicative of response to the linguistic stimulus (Fernald et al., 2008). As children heard two iterations of the test sentence, we form two corresponding windows of analysis for 3000 ms after the disambiguation point in each iteration of the sentence. In all cases, the disambiguation point has been shifted forward in time 300 ms, to account for time it takes young children to plan an eye saccade, ensuring that all responses are responses to the target audio (Fernald et al., 2008). Trials where children were attending to the screen less than 60% of the time were eliminated from the analysis. To analyze the data I used empirical logit mixed effects models, fit in R (R Core Team, 2013) with the *lmer()* function of the *lme4* library (Bates, Maechler & Bolker, 2013). The dependent measure in the models was the empirical logit transform of

the proportion of time spent looking to the other-directed action. Variables included in the analysis were fixed effects of participants' age and MCDI vocabulary score, as well as fixed effects of window (first or second iteration of the test sentence) and condition (NAME condition as in (29) and PRONOUN condition as in (30)), and all interactions, and random effects for participant and item.

As noted above, the three different interpretive strategies identified here predict three distinct response patterns. Knowledge of Principle C predicts a preference for the other-directed action in response to (29) but no preference for either image in response to (30). This would surface as a main effect of condition. Both interpretive strategies based in a transitivity bias predict the behavior to be stable across condition, which would show as a lack of a main effect of condition. The Event-level Transitivity Bias predicts a preference for the other-directed action in both cases, corresponding performance significantly higher than chance in both conditions. The Sentence-level Transitivity Bias predicts no preference for either event in both cases, corresponding to chance performance in both conditions.

2.3.2 Results

While the preliminary model included participants' age and MCDI vocabulary score, I compared this model to those where these terms were absent in order to determine whether they contributed to model fit. Model fit was not significantly improved by including either Age ($\chi^2(8)=9.83$, $p=.277$) or MCDI Vocabulary score ($\chi^2(8)=13.79$, $p=.087$), and these terms were subsequently dropped from the remainders of the analyses.

Figure 3 shows the mean proportion of looking to the other-directed event across the timecourse of the trial. Windows of analysis are identified by the outlined boxes super-imposed over the graph. Prior to hearing the disambiguating object NP, children do not look preferentially to either event. However, during the critical windows following the object NP, children looked significantly more to the other-directed action in the NAME condition, while showing no preference in the PRONOUN condition. This performance is maintained across both windows of analysis. There is a significant main effect of Condition ($\chi^2(1)=4.92$, $p<0.05$), no effect of Window ($\chi^2(1)=0.232$, $p=0.63$), and no interaction ($\chi^2(1)=0.094$, $p=0.759$).

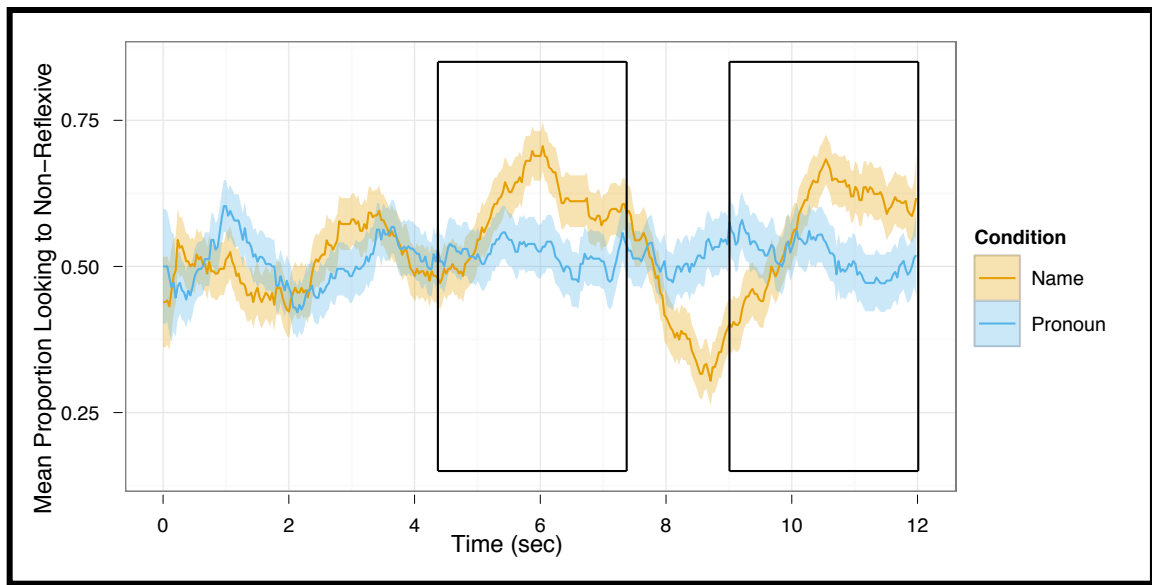


Figure 3 Looking behavior in Experiment 1. Outlined boxes represent critical windows for analysis (3000ms following the onset of the disambiguating object NP in two iterations of the test sentence).

2.3.3 Discussion

The results from Experiment 1 showed that at 30 months, children's interpretations of sentences like (29) and (30), repeated here as (35) and (36), show a consistent asymmetry. Children prefer to interpret (35) as referring to an event

corresponding to a disjoint interpretation (an other-directed action), while they have no preference for interpreting (36) as disjoint or coreferential. This behavioral pattern cannot be accounted for with either formulations of a transitivity bias.

(35) She's patting Katie's head.

(36) She's patting her head.

One additional finding worth consideration at this time is the non-effect of Vocabulary size. In the comparison of model fit, MCDI vocabulary size was shown not to improve model fit, meaning that it did not contribute to explaining variation between subjects in this task. Interestingly, this is a non-replication of the findings of LCL, who demonstrated that children with higher vocabulary scores showed stronger preferences for a disjoint interpretation in Principle C contexts. The results presented here, however, show no significant mediation by vocabulary size. This divergence from the findings of LCL could stem from several possible factors. First, it could be the result of less variation in Vocabulary size in our sample than in LCL's. The mean Vocabulary score in our sample was 560 (of 680) compared to 446 as reported by LCL. One may note that this age range is at (and extending beyond) the upper limit of the age range on which the MCDI is used. In effect, this measure may be hitting a ceiling. It is possible that we may see an effect of vocabulary size if we had another way to measure it; something more extensive than the MCDI. Future research could benefit from further exploration of vocabulary measures and their connection to grammatical competence.

Alternatively, our findings may differ from LCL with respect to a Vocabulary effect due to methodological differences, in two ways. It could be that the Vocabulary effect observed by LCL was driven by their analysis of the data over a large time window; recall that Lukyanenko et al. compared performance before and after hearing the first repetition of the test sentence. The post-utterance window was 9 seconds long. It may be that a similar pattern is present in their data, and simply gets washed out in such a large window of time. That is, it could be that even low vocabulary children could be shown to succeed on LCL's task in a smaller time window but this effect is masked over the long window of analysis. This non-replication could also arise due to differing subject designs: while LCL used a within-subjects design, meaning that children heard four trials with each sentence type, I used a between-subjects design, where children heard eight trials all with one sentence type. Having to interpret multiple sentence types in LCL's could have obscured some children's ability to demonstrate their knowledge, resulting in a difference in performance across Vocabulary score. I further explore vocabulary size and its possible effects on the acquisition of Principle C in Chapter 4.

Experiment 1 does replicate LCL's primary finding of observing so-called 'Principle C effects'- preferences for a disjoint interpretation in Principle C contexts- at 30 months. Further, comparison to similar transitive sentences which are not subject to Principle C showed that this behavior is in fact more consistent with interpretation via Principle C than interpretation via a bias found in interpreting transitive syntax. However, the question remains as to whether this is the youngest age at which this restriction on interpretations might be observed. In Experiment 2, I explore children's performance at

24 months in the same contexts as that of Experiment 1 in order to identify the lower bound on observable Principle C effect behavior.

2.4 Experiment 2

To my knowledge, as of this writing, the findings of LCL and those presented in this dissertation provide the only data points for comprehension in Principle C contexts at 30 months, and no known studies have probed any younger. LCL argued that low-vocabulary children in their sample failing to exhibit a restriction in interpretation in these contexts was potentially suggestive that this age marks the onset of Principle C effect behavior. However, given that we find no mediation of vocabulary size in Experiment 1 (a non-replication of LCL's result), I consider here the possibility that Principle C effects could be observed even younger than 30 months.

How likely are children at 24 months to be able to perform in this task? There are a few aspects of the task that we may consider here. As a method of analyzing children's interpretations, preferential looking is used as young as 2 months of age (Kuhl & Meltzoff, 1982; Baier, Idsardi & Lidz, 2007). In terms of the syntax of the test sentences, children have been shown to recognize sentences with similar transitive structures in preferential looking contexts as early as 21 months (Gertner, Fisher & Eisengart, 2006). Possessive inflection, relevant for comprehension of the possessive object NP, is among the earliest morphemes in children's acquisition (Brown, 1973), and while this inflection does not emerge in children's speech until around 3 years old, children as young as 2 years have been reported to produce possessive structures with the inflectional morphology omitted, as in (37) (Radford & Galasso, 1998).

- (37) a. “That Mommy car.”
b. “It Daddy bike.”

As comprehension has been robustly shown to outstrip production in language acquisition, we can be reasonably certain that possessive nominal structures should pose no problem for children’s comprehension by 24 months.

Although it seems likely that children will be able to form an interpretation of our test sentences by 24 months, how likely is it that children will exhibit a restriction on interpretation at this age? Given that transitive syntax is a strong cue to interpretation in word-learning contexts well earlier than the second birthday (as young as 19 months), it is also likely that children have at least one possible mechanism for narrowing interpretations in this context; even if this age precedes the onset of Principle C, we could still find evidence of Principle C effects driven by a transitivity bias. While Experiment 1 demonstrated that behavior at 30 months is not consistent with a transitivity bias interpretive strategy, this does not preclude the possibility that such a strategy could be the primary mechanism for interpretation at an earlier stage in development. Experiment 2 represents the first foray into children’s interpretations in Principle C contexts at 24 months.

2.4.1 Method

48 24 month-olds (24 males) with a mean age of 24;10 (range 23;3 to 24;24) were included in the final sample. An additional 2 children were excluded due to equipment

failure during testing. All materials, design, procedure, and analysis were identical to Experiment 1.

2.4.2 Results

As in Experiment 1, Age and MCDI Vocabulary score were included in a preliminary model, then this model was compared to those where these terms were absent in order to determine whether they contributed to model fit. Model fit was not significantly improved by including either Age ($\chi^2(8)=2.69$, $p=.953$) or MCDI Vocabulary score ($\chi^2(8)=11.17$, $p=.192$), and these terms were subsequently dropped from the remainders of the analyses.

Figure 4 shows the mean proportion of looking to the other-directed event across the timecourse of the trial. Windows of analysis are identified by the outlined boxes super-imposed over the graph. Children do not have a significant preference for either event prior to hearing the disambiguating object NP; however, unlike 30 month-olds, the disambiguating point does not provoke a strong preference for either interpretation in either condition. Looking behavior oscillates near chance throughout the trial, with a slight increase in attention to the self-directed event preceding the first window of analysis, and a preference for the other-directed event during the first window. This preference for the other-directed event in the first window is realized as a significant main effect of Window ($\chi^2(1)=6.779$, $p<0.01$), no effect of Condition ($\chi^2(1)=0.017$, $p=0.898$), and no interaction ($\chi^2(1)=1.768$, $p=0.184$). Planned comparisons reveal that looking in the pronoun condition is significantly different than chance ($t(22)=3.392$, $p<0.01$); children prefer the other-directed action for sentences like (23) upon first

hearing this sentence. No other looking patterns in the critical windows differ significantly from chance.

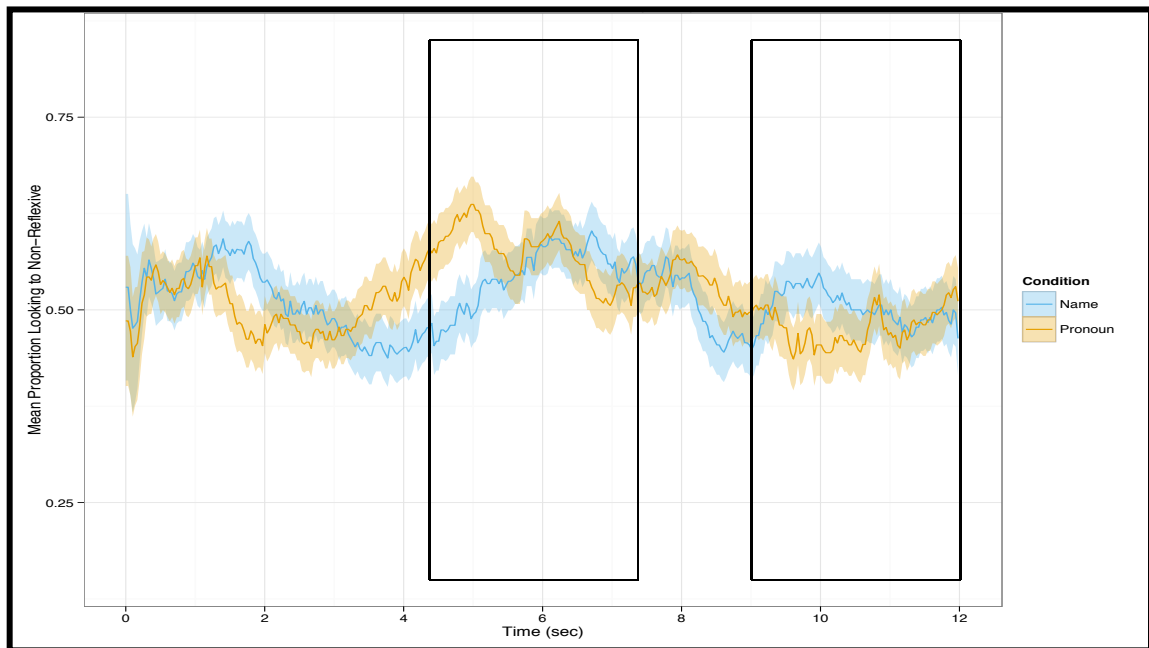


Figure 4 Looking behavior in Experiment 2. Outlined boxes represent critical windows for analysis (3000ms following the onset of the disambiguating object NP in two iterations of the test sentence).

2.4.3 Discussion

Results from Experiment 2 suggest that children do not exhibit classic Principle C effects at 24 months. While children do show a preference for the other-directed action, corresponding to a disjoint interpretation, following the first test sentence iteration, this effect is carried by performance in the PRONOUN condition. Performance in the NAME condition, where Principle C effects would be most expected, is not different from chance. This behavioral pattern is not predicted by Principle C, nor either of the versions of a transitivity bias specified here. Upon examining the graph more closely, it is evident that the looking patterns responsible for this initial preference in the pronoun condition

arises well before the critical window (approximately near the beginning of the first test sentence). Recall that this window begins at the onset of the object NP, i.e. the point at which the sentence is disambiguated between two possible interpretations. Thus it seems like the sole effect observed here cannot be directly attributed as a response to the linguistic stimuli of the task. From this result we can conclude that children's performance is not consistent with either knowledge of Principle C or with a strong preference driven by a bias in transitive structures.

2.5 General Discussion

Two experiments demonstrated that the earliest restrictions on interpretations in Principle C contexts are in place by 30 months, and further that these restrictions are more consistent with accurate knowledge of Principle C than with an interpretive bias in transitive clauses. In fact, barring the unlikely possibility that children rapidly move through an intermediary stage in the three months between the 23-25 month and the 28-32 month ranges tested in these two experiments, it seems to be the case that children never utilize this interpretive strategy to resolve reference, although it is robustly utilized in verb learning contexts..

We might consider this possibly surprising non-finding at this point: for all the robust findings on the bias to treat transitive structures as involving two participants, why don't we find evidence of this? Given the strength of transitive syntax as a cue to particular meanings, as evidenced by myriad syntactic bootstrapping studies, one might be surprised that such cues do not seem to be used here. If these biases are so pervasive, why are they not shown to be utilized in our task?

For the case of 30 month-olds, it seems likely that children have access to a better interpretive mechanism. One possibility is that by this age children have successfully acquired and put into use accurate knowledge of Principle C. However, this is not the only possibility; in Chapter 3, we will discuss another alternative interpretive strategy that is consistent with all behavior observed thus far.

Irrespective of what interpretive mechanism 30 month-olds *are* using, they are likely *not* using a bias based in verb transitivity because although they share the same syntax, this is a different type of interpretive uncertainty than what children face in bootstrapping tasks. In tasks where transitivity has been shown to be a successful cue to interpretation, such as bootstrapping tasks, the uncertainty children face is in the interpretation of the verb. In Principle C contexts, however, the uncertainty is related to the interpretation of reference relations- an entirely different task. In this context, transitivity is not a reliable cue to interpretation and thus children are unlikely to implement this interpretive strategy. This reasoning could also account for the behavioral pattern at 24 months, where performance is largely chance-like. Even in the absence of another consistent strategy, children do not implement a transitivity bias because they recognize that it is not the applicable context.

2.6 Conclusion

Children have been shown to utilize argument structure of sentences as a strong cue to the types of interpretations those sentences will have. These findings suggest a possible mechanism for interpretation that children could use to interpret simple transitive sentences in Principle C contexts. In simple sentences and contexts like those

tested by Lukyanenko, Conroy & Lidz (2014), such a transitivity bias predicts the same behavior as knowledge of Principle C.

In this chapter, I sought to answer two questions. First, what is the lower bound for observing Principle C effects in children? Is 30 months really the onset of Principle C behavior? Second, at the youngest age at which children exhibit Principle C effects, what underlying knowledge drives this behavior? Is performance more consistent with knowledge of Principle C, or a bias to treat transitive structures as always involving two participants?

In Experiment 2, I demonstrated that 24 month-olds do not consistently exhibit Principle C effects- behavior in Principle C contexts is not different from chance. This finding serves to answer our first question: the onset of behavior consistent with knowledge of Principle C occurs between 24 and 30 months of age. In Experiment 1, I found that across multiple contexts, 30 month-olds interpretations are more consistent with knowledge of Principle C than with a transitivity bias. Neither formulation of a bias in interpreting transitive structures is able to account for children's responses to all sentence types. Together, these results suggest that even at the earliest point when restricted interpretations in Principle C contexts are observed, children never utilize a heuristic based in transitivity to interpret sentences in Principle C contexts.

3 A Bias Against Backward Anaphora is Not Sufficient

3.1 Outline

Chapter 3 examines the possibility that children's early interpretations in Principle C contexts could be attributed to an interpretive strategy which favors a disjoint interpretation in all cases of backward anaphoric structures, rather than relying on structural knowledge of Principle C. Research in children's early interpretation of sentences has suggested that children utilize linear order as part of a reduced representation before they are able to access a full structural representation. This suggests that children could potentially hypothesize an initial grammatical rule to account for the pattern of reference relations by appealing to linear order. Research in child and adult interpretations of backward anaphoric structures have also shown them to be subject to discourse conditions. When the appropriate context is not met, a disjoint interpretation is preferred for all cases regardless of structure. This suggests children could utilize a simple general preference for disjoint interpretations in all cases of backward anaphora. Both of these possible linear mechanisms provide an alternative to Principle C by which children could show a preference for disjoint interpretations in simple Principle C contexts.

This chapter provides a novel exploration of 30 month-olds interpretations of backward anaphoric structures where a pronoun precedes but does not c-command an R-expression, and thus a coreferential interpretation is grammatically available. Comparison to interpretations of minimally different sentences which *are* constrained by Principle C

effects shows that all backward anaphoric interpretations are not blocked at 30 months. Findings suggest that an overall bias for disjoint interpretation of backward anaphoric structures cannot account for interpretations across these contexts.

3.2 Introduction

Children have been shown to interpret simple sentences like (38) similarly to adults- disallowing a coreferential interpretation- by 30 months of age.

(38) She's patting Katie.

While Lukyanenko, Conroy & Lidz (2014) suggest that this result is likely indicative of children having in place a structural constraint against coreference in these cases, it is important to remember that this is not the only explanation for such behavior. Although this behavior pattern is consistent with children using knowledge of Principle C to form interpretations, this alone is not sufficient to attribute knowledge of Principle C. A preference for a disjoint interpretation for sentences like (38) is equally consistent with numerous other mechanisms that could drive identical behavior.

Chapter 2 examined one such potential mechanism- a bias to interpret arguments in transitive clauses as disjoint in reference. The results of Experiment 1 suggested that such a bias could not account for behavior across all contexts of 30 month-olds' performance. This chapter examines another possible interpretive strategy: a bias to interpret all backward anaphoric structures as disjoint in reference. In this chapter I review evidence from child and adult interpretations of backward anaphoric structures,

showing that an overall bias is exhibited even in ambiguous contexts, where a coreferential interpretation is grammatically available. In simple single-clause structures such as those tested by LCL, such a bias predicts the same performance as knowledge of Principle C. These findings suggest that a bias for backward anaphoric structures could account for the preference for a disjoint interpretation for (38) exhibited by 30 month-olds.

The research presented here seeks to examine the nature of the knowledge state driving interpretations of sentences like (38) at 30 months. I compare performance predicted by adult-like grammatical knowledge of Principle C and an alternative bias to interpret all backward anaphoric sentences as disjoint. These two possible mechanisms for interpretation predict identical performance for simple sentences like (38), used in LCL's task and Experiments 1 and 2. Here, I utilize the context of backward anaphoric structures where c-command relations between a pronoun and R-expression vary in order to identify which mechanism is responsible for children's interpretations. I examine interpretations of two backward anaphoric structures, identified in sentences (39) and (40) below. Because the linear order of the nominal elements is maintained in both sentences, a bias to interpret all backward anaphoric sentences as disjoint in reference predicts the same performance for each sentence. However, because binding obtains in (39) but not (40), Principle C predicts a restriction on interpretation only in one case.

(39) She's painting the house that's in Katie's lap.

(40) The house that she's painting is in Katie's lap.

In the current study I compare children's responses to sentences like (39) and (40). An asymmetry in the available interpretations is expected by Principle C but not by a broader bias applying to all backward anaphoric sentences. I show that children's response patterns to such sentence types differ; this finding suggests that a broad bias for all backward anaphoric sentences cannot account for children's behavior, and is thus not likely to be the mechanism driving children's interpretations. In the remainder of this introduction, I recognize findings suggesting that linear order of string elements is a pervasive cue that children recognize from early ages, and identify observations from child and adult research on backward anaphoric structures suggesting an overall bias for interpretation exists. I then present an overview of the current study which was designed to test 30 month-olds' interpretations in similar backward anaphoric contexts.

Linear order and directionality effects in early interpretation

Linear order of elements in a string is an inherently prevalent source of information; the modality (or modalities, when considering signed language systems) in which language is expressed requires that only one unit can be expressed at a time, and thus linguistic elements are inherently ordered in time with respect to one another.

Knowledge of word order and its effect on the meaning of sentences has been shown to be present in children as young as 16 months. Golinkoff et al. (1987) presented children with images of events matching the scenes described in (41) and (42). They found that by 16 months, children attend more to the event that matches the description they hear, suggesting that they recognize the links between word order and thematic interpretations in their language.

(41) Cookie Monster is tickling Big Bird.

(42) Big Bird is tickling Cookie Monster.

Linear order of words in a string is clearly a cue that children have access to from very early stages of development. Further, it seems that children may rely on linear order for interpretation before they have access to accurate representations of the structure of sentences. Gertner & Fisher (2012) found that 21 month-olds begin to interpret sentences by assigning thematic roles based on the order in which nouns appear. They presented children with conjoined-subject intransitive sentences such as (43), along with events that depicted an asymmetric action (e.g. boy spinning girl in a chair) and a simultaneous action (e.g. boy and girl both waving wands).

(43) The boy and the girl are gorpig!

Recall from our discussion in Chapter 2 that by 24 months, children interpret such sentences as reflecting a simultaneous action (Naigles, 1990). Gertner & Fisher, however, found that at 21 months, children interpret sentences like (43) as corresponding to an asymmetric action. Their results suggest that children initially rely on partial representations utilizing the number and order of nouns to assign thematic roles. Thus it seems that linear order is a potential source of interpretation, and could be used in children's early interpretation of reference relations.

Many of the cases where a pronoun precedes its potential antecedent are also those where it c-commands it and Principle C applies to restrict interpretation. Because we know that children's initial interpretations rely on linear order, it is possible that children could develop an initial strategy for interpreting reference relations that relies on linear order rather than structural restrictions like Principle C. While this bias only approximates Principle C and cannot account for interpretations in all cases, however linear order and c-command in large part co-occur. Thus one possibility for attributing early Principle C effects to a linear order interpretive strategy is that children have encoded a grammatical rule based in linear order as an initial generalization before identifying the correct rule.

A second possibility as a way that linear order could affect interpretation in 30 month-olds is suggested by the findings that backward anaphoric structures are subject to strict discourse conditions. Consider the set of sentences in (44-47).

- (44) While Katie₁ was in the kitchen, she_{1/2} baked cookies.
- (45) While she_{1/2} was in the kitchen, Katie₂ baked cookies.
- (46) Katie₁ baked cookies while she_{1/2} was in the kitchen.
- (47) She_{1/*2} baked cookies while Katie₂ was in the kitchen.

The pronoun *she* can derive its reference either anaphorically by relating to another nominal element, or deictically by referring directly to an individual. In sentences such as (44) and (45), the pronoun follows the R-expression antecedent *Katie*; this directional relation is referred to as anaphoric or forward anaphoric. In sentences such as

(46) and (47), the pronoun precedes the R-expression; this relation is referred to as cataphoric or backward anaphoric. Binding constraints only place a restriction on interpretation in the case of (47), where *she* c-commands the R-expression *Katie* and thus Principle C disallows a coreferential interpretation.

By age 3, children have been shown to accept coreferential interpretations for backward anaphoric sentences like (46) significantly more often than they do for sentences like (47), where a coreferential interpretation would be grammatically unavailable for adults. Many studies additionally report a preference for a disjoint interpretation for all cases of backward anaphora, including those where a coreferential interpretation is grammatically available (see Lust, Eisele & Masuka 1993 for a review of these findings). However, this effect may be the result of the discourse conditions for a backward anaphoric (i.e. coreferential) interpretation not being appropriately satisfied in these tasks. For example, Lust, Loveland & Kornet (1980) explored children's comprehension of sentences similar to (45), and report an average of 29% coreferential interpretations. However, when they included a pragmatic lead which introduces the potential antecedent as a discourse element, coreferential interpretations jumped to 59%. Studies that have reported similarly low rates of coreferential interpretations (e.g. Eisele (1988); Eisele & Lust (1990)) have not included such introductory sentences. This failure to satisfy discourse conditions could explain the preference for disjoint interpretations over coreferential interpretations.

This preference for a disjoint interpretation in backward anaphoric sentences where coreference is grammatically available has also been attested in adults. Gordon & Hendrick (1997) showed that adults reject a coreferential interpretation for single clause

backward anaphoric sentences like both (48) and (49), irrespective of which interpretation(s) the syntax makes available. Again, however, this effect was only pervasive when no context was supplied; availability of the coreferential interpretation was shown for (48) to increase when context was provided.

(48) His roommates met John at the restaurant.

(49) He met John's roommates at the restaurant.

These findings suggest another way in which linear order could affect interpretation in young children. Rather than a grammatically specified rule as an initial interpretive mechanism, children could alternatively have a preference for disjoint interpretations which is derived from the same source that makes backward anaphoric sentences subject to strict discourse contexts in adults.

There are therefore at least two ways in which linear order could be used as a cue for interpretation instead of knowledge of Principle C. However, both a grammatical rule encoded as an initial interpretive strategy and a pragmatically-induced bias predict exactly the same behavior: preference for a disjoint interpretation whenever a pronoun precedes an R-expression. Because we cannot distinguish between behavior caused by one or the other, I will collapse across these possibilities for the remainder of this discussion and refer to the effects of both as coming from a 'Linear Bias' (defined below).

The Linear Bias poses a potential challenge for the conclusion that LCL's findings are indicative of knowledge of Principle C at 30 months. Because the simple

sentences tested by LCL are exactly those in which a bias against backward anaphora would apply, it is possible that the behavior patterns exhibited by 30 month-olds could be a more general response to backward anaphoric sentences, rather than adult-like interpretation based in structural constraints on reference relations.

The present study

In the present study, I address the question of whether a disjoint interpretation for sentences in Principle C contexts could be the result of an alternative mechanism for interpreting all backward anaphoric structures as disjoint in reference. We can formalize such a bias by appealing to linear order between an R-expression and other nominal elements within a sentence, as in (50):

(50) Linear Bias: a pronoun cannot precede its antecedent.

Because linear order and c-command relations are confounded in English, this Linear Bias will predict the same disjoint interpretation as adult grammatical knowledge would in all cases where Principle C applies, including sentences like (38) tested by LCL. In order to identify the mechanism that drives interpretation at 30 months, it is necessary to identify contexts in which these mechanisms would predict differing performance. The critical distinction between a Linear Bias and knowledge of Principle C are cases where a pronoun linearly precedes but does not c-command an R-expression; in such cases, Principle C does not apply and thus does not restrict interpretation. However, because children and adults have been shown to exhibit a preference for disjoint interpretations

even in these cases where the coreferential interpretation is grammatically available, it will be necessary to compare performance across minimally different sentence types, where c-command does and does not obtain, to determine whether performance varies with binding conditions. While Principle C predicts a stronger preference for the disjoint interpretation in cases where c-command of a pronoun over an R-expression obtains, a Linear Bias predicts performance to be uniform across all cases of backward anaphoric structures.

Adjunct clauses such as those in (44-47) above can be used to vary c-command relations while maintaining linear order between nominal expressions, and indeed these structures are commonly used to test the interpretations available to children (and adults), with methodologies such as the Truth Value Judgment Task. Experiments employing this methodology typically present a subject with a short story, following which a description of the story is given; the subject's task is to identify whether the sentence is an accurate description of the story. To test whether the coreferential interpretation is made available by the grammar for a given structure, the disjoint interpretation of the test sentence is made false in the context of the story, and the coreferential interpretation is true by the context. The TVJT design is based on the claim that if all else is equal, subjects will obey the 'Principle of Charity,' assenting to the truth of a sentence if they can, that is, if any possible interpretation is both available by the grammar and true by the context⁷ (Crain & Thornton, 1998). Thus if children judge the target sentence to be false, this can be taken

⁷ However, rarely is all else equal- the preference to respond to an interpretation made true by the context is shown to have been overridden in a number of contexts (Gualmini et al., 2008; Conroy et al., 2009; Viau, Lidz & Musolino, 2010; Syrett & Lidz, 2011).

as evidence that the interpretation which is true by the context (the coreferential interpretation) is not available in the child's grammar.

While adjunct structures are thus ideal for testing the application of constraints with such TVJ tasks, this methodology is not effective for testing children at 30 months, due to the extra-linguistic requirements of such a task. In order to respond accurately during test, children must remember a series of actions which took place in the story, which may be beyond the memory and recall capabilities of children this young. Additionally, the requirement to make an overt choice or response can be cognitively demanding and has the potential to obscure underlying understanding (Hamburger & Crain, 1982). Because of the complexity of the scenes that would be required to represent the interpretations of such sentences, adjunct clauses are less than ideal for testing in a Preferential Looking paradigm.

Take, for example, the set of events which would be needed to represent the coreferential and the disjoint interpretation of a sentence such as (51) or (52). Each event would require two characters- Susan and another girl- and two actions- eating breakfast and reading a book. While Susan reads a book in both scenes, in one event, the breakfast-eating would be done by Susan, and in the other it would be done by the second girl. Such scenes would inherently need a large amount of information to be conveyed in order to appropriately represent the sentence interpretations. Representing this all in two simultaneously-presented scenes could easily be too overwhelming for 30 month-olds to be able to appropriately map their interpretation of the sentence to the visual scene.

(51) While she ate breakfast, Susan read a book.

(52) She ate breakfast while Susan read a book.

Another structure that allows for the manipulation of c-command relations is possessive structures, as in (48) and (49) above. However, these structures are not ideal for our purposes for two reasons. First, single-clause backward anaphoric structures such as these are those that have been shown to be most affected by discourse context; these cases require very specific discourse conditions to be licensed. Additionally, they have the challenge of not being representable across the sentence types by the same two events. Consider (53) and (54) below; the two potential interpretations of (53) correspond to events where Susan's mother or another girl's mother hugs Susan. The potential interpretations of (54) correspond to events where Susan or another girl hug Susan's mother.

(53) Her mother hugged Susan.

(54) She hugged Susan's mother.

An ideal case for comparing children's interpretations in backward anaphoric sentences within a Preferential Looking Paradigm would contain a minimal number of actions to be represented in the visual scenes, as well as allowing comparison of interpretations of the two sentence types to be diagnosed with the same set of two events. In the present study, I compare performance to sentences with relative clauses, as in (55) and (56) below.

(55) She's painting the house that's in Katie's lap.

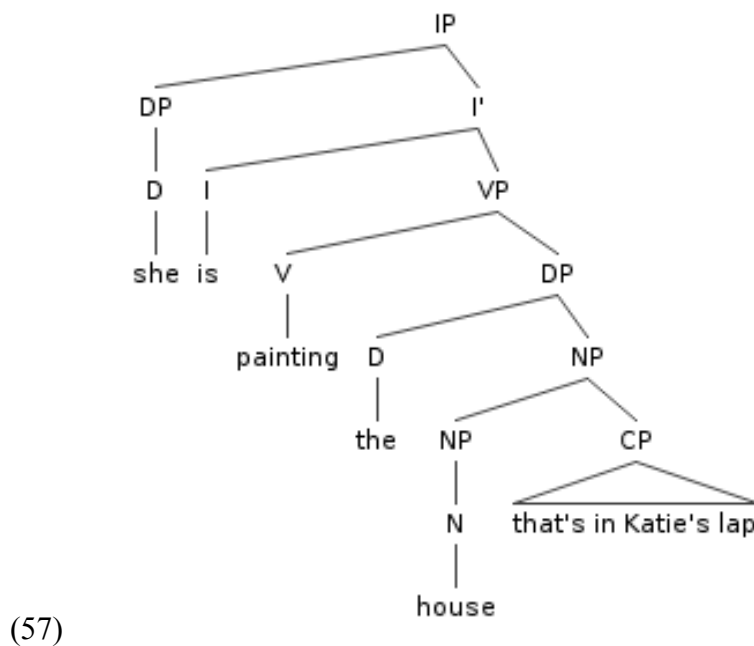
(56) The house that she's painting is in Katie's lap.

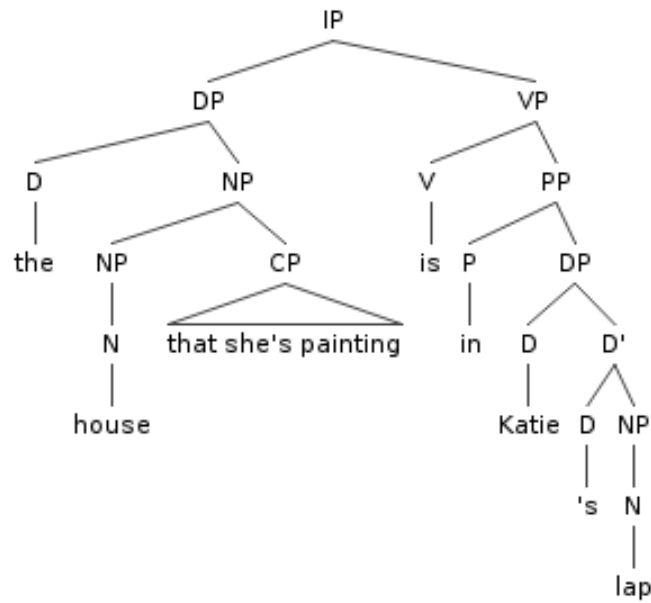
In requiring only one action to be performed and in allowing the disjoint and coreferential interpretations of each sentence type to be represented by the same events, these structures present an ideal case for testing in a Preferential Looking Paradigm.

Like adjuncts and possessive structures, these relative clause structures also allow the manipulation of c-command relations while maintaining linear order. Because only one of the two interpretive mechanisms under consideration here takes c-command relations into account, the response pattern across these two types of sentences should allow us to identify whether c-command relations affect interpretation, identifying which of these mechanisms is driving interpretation. I will now consider each of these interpretive strategies individually to lay out the predictions that it makes for each sentence.

Given the structure in (57) for the string in (55), the R-expression *Katie* is contained within the c-command domain of the pronoun *she*. As such, Principle C disallows co-indexing of *she* and *Katie*, blocking the coreferential interpretation for (55). However, given the structure in (58) for (56), c-command does not obtain in the same manner. Because the pronoun *she* is embedded within the relative clause, its c-command domain is limited to [*I' is painting*]. Thus even if *she* and *Katie* are co-indexed, no binding relation could hold between them and Principle C would be satisfied. (56) can have either a coreferential or a disjoint interpretation. Therefore if Principle C is the mechanism which drives children's interpretation of such sentences at 30 months, we

predict a preference for a disjoint interpretation of (55), and no grammatical restriction of interpretation for (56). However, recall that previous research has found an overall discourse preference for a disjoint interpretation in backward anaphoric sentences. Because of this, it is possible that children may exhibit such a preference for both sentence types. To avert this effect, the present study will be designed to bias as much as possible toward a coreferential interpretation (while still allowing the disjoint interpretation to be plausible), so that if a disjoint preference for either sentence type is observed, it is because the coreferential interpretation is grammatically unavailable. Methods for creating this coreferential interpretation bias are discussed in the Design section 3.3.1.





(58)

Notice however that the pronoun linearly precedes the R-expression in both sentence types. A Linear Bias as stated in (50) predicts a lack of a coreferential interpretation for both cases. Thus if interpretations are driven by a Linear Bias at 30 months, we predict a preference for the disjoint interpretation to be exhibited uniformly across both sentence types.

Experiment 3 examines 30 month-olds' performance in a preferential looking paradigm which presented arrays corresponding to the coreferential and disjoint interpretations of sentences like (55) and (56). I compared attention following the test sentence to determine whether children exhibited a preference for the disjoint interpretation in response to either sentence type. The results demonstrate that 30 month-old children exhibit a preference for the disjoint interpretation in response to sentences like (55) but not (56); this performance is expected only if Principle C is the source of

knowledge driving interpretation, and cannot be accounted for by appealing only to a linear bias.

3.3 Experiment 3

3.3.1 Method

Participants

52 30 month-olds (26 males) with a mean age of 30;10 (range 28;6 to 31;27) were included in the final sample. An additional 15 children were excluded from the final sample for the following reasons: equipment failure/ experimenter error (n=12), failure to complete the task (n=1), failure to meet language input criterion (n=2). Participants were recruited through the University of Maryland Infant and Child Studies Database from the College Park, MD area. Children participated on a volunteer basis, and were given a small gift for participating. All participants were learning English at home, and English input constituted at least 80% of their language input.

Materials & Design

Visual stimuli were designed similarly to those presented in Experiments 1 and 2. All video was presented on a 51" plasma TV screen. Videos were edited together to create the sequence of events outlined in the Design section below. A full schematic of the events children see throughout the video can be found in Appendix B. Single videos were presented in the center of the screen and were sized 18 inches wide and 29 inches tall. Videos presented together in preferential looking format were always identical in

size and scale, measuring 16 inches wide and 11 inches tall, with a 10.5 inch span between the inside edges of the two images.

Audio stimuli was recorded in a sound-proofed room by a female native speaker of American English using intonation common in Child Directed Speech. Recordings were edited and combined with the visual stimuli.

The components of the video mirrored those of the task utilized by LCL and presented in Experiments 1 and 2. A Character Introduction phase first familiarized children to the two characters through short video clips (3.5-6.5 seconds each), which named each character (Anna and Katie) individually and described them doing simple actions. In addition to occurring at the beginning of the sequence, several Character Introduction clips were placed throughout the video to provide filler material between test trials as well as to further remind children of the character names. Following the Character Introduction phase, children received six Face Check sequences, which tested children's memory of the character names to facilitate comprehension during test, as well as preparing children for the preferential looking format of the test trials. During the Face Check sequences (6 seconds each), the two characters were presented in separate videos simultaneously in a preferential looking paradigm while a sentence asked children to find one of the two characters (e.g. *Where's Katie? Do you see Katie?*). Target character and the side of the screen on which the target character appeared were counterbalanced across Face Check trials. Order of Face Check trials was counterbalanced across subjects.

Figure 5 presents a schematic of a typical Familiarization and Test phase sequence. Events depicted in the familiarization and test trials were all continuous or

repeatable two participant events in which one character performed an action on an object.

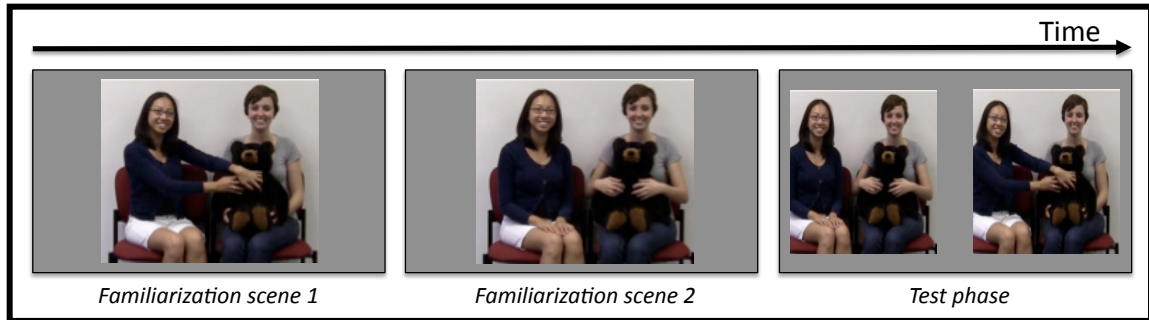


Figure 5 Schematic of Familiarization and Test phase sequence in Experiment 3.

Actions were chosen to correspond to transitive verbs that are the most commonly produced by children at 30 months. All verbs were reported to be produced by at least 70% of children at this age (with the majority produced by over 90% of children). Objects used as the patient of the event were common household objects and animal puppets, the names for which are reported to be used by 85% or more of children at 30 months. The scenes for a given test trial consisted of two versions of an event in which each character acted upon the object. For both scenes, the object was placed in the lap or the hand of one of the characters- this location remained constant across both scenes. In each scene, one of the two characters performed an action on the object. Both characters were always present in both scenes, even if one of the characters did not participate in the action or in holding the object.

Immediately preceding each test trial, children were presented with a Familiarization phase. In two video clips (6 seconds each), the test events were presented one at a time. Audio described the event by identifying which character was performing the action, as in (59).

(59) Wow! There's Anna and Katie! Anna's painting a house!

Oh look! There they are again! Now Katie's painting the house!

Because children (and adults) show an overall bias toward a disjoint interpretation of ambiguous backward anaphoric structures, we aimed to make the coreferential interpretation as available and prominent as possible (while still making the disjoint interpretation a plausible interpretation). This coreferential bias was included so that if a preference for the disjoint interpretation was observed, it would be more likely to be caused by a constraint against a coreferential interpretation exhibited in the grammar, rather than a more general preference for interpretation. To this end, the event which corresponded to the coreferential interpretation was always shown second, immediately preceding the test trial. This served to make this event more prominent, as well as making the character named in this familiarization more accessible in the discourse, promoting interpretation of any following pronouns as referring to this character if possible.

The Test phase (12 seconds each) immediately followed with the same two event videos as shown in the Familiarization phase presented side by side simultaneously. An attention-getting filler phrase (*What's happening now?*) introduced the events. Test audio consisted of two repetitions of the test sentence. Children were presented with a total of 8 test trials in a between subjects design: half of the subjects heard MATRIX PRONOUN condition sentences as in (60) for all trials, and half heard all EMBEDDED PRONOUN condition sentences as in (61). Trial order and the side of the screen each video appeared on was counterbalanced across subjects.

(60) She's painting the house that's in Katie's lap!

(61) The house that she's painting is in Katie's lap!

In addition to the phases described above, children also saw 4 filler video clips (7 seconds each) spaced throughout the video. These clips depicted children's toys presented with classical music, and were used to attract children's attention to the screen, as well as to provide a 'mental break' between the more challenging test trial sequences. For all video clips, there was 1 second black screen between each clip, and the first audio clip began 20 frames (.67 seconds) after the visual stimuli appeared on the screen.

Procedure

Procedure for running subjects was identical to that of Experiments 1 and 2.

Analyses & Predictions

We expect children's response pattern to differ substantially in terms of the timecourse of response as compared to that observed by LCL or in Experiment 1 for two reasons. The first is that the sentence is much more complex than the single clause structures used in those studies. Because there is significantly more structure to parse through⁸, arriving at an interpretation may take longer. The second reason we expect a response delay is that the visual array is also more complex; whereas LCL and

⁸ This is the case even if children are using a Linear Bias to interpret the reference relations within the sentence- they still need to use some form of a parsing strategy to interpret the remainder of the structure of the sentence to arrive at an interpretation.

Experiment 1 had two images which each included two characters and an action, this study incorporates an additional object into each scene. Because there is more visual information to keep track of, children may need to perform more ‘check-back’ looks, where they re-orient to a distractor image to confirm that it does not match their interpretation.

For these reasons, the window of analysis in this experiment is identified as the first 3000ms following the onset of the second repetition of the test sentence⁹. This should allow enough time for children to work through the structure and also the visual array; additionally, the onset of the second test sentence iteration may focus children’s attention on their chosen interpretation.

The last trial was excluded from analysis for failure to meet the criterion for attentiveness (85%). To analyze the data I used empirical logit mixed effects models, fit in R (R Core Team, 2013) with the *lmer()* function of the *lme4* library (Bates, Maechler & Bolker, 2013). The dependent measure in the models was the empirical logit transform of the proportion of time spent looking to the disjoint interpretation action (defined as the event where the character holding onto the object was not the one acting on the object). As noted above, performance will be examined for a main effect of condition, to signal a difference in interpretation across these sentence types. Knowledge of Principle C predicts a preference for a disjoint interpretation in response to (60) but not (61), whereas a Linear Bias predicts a preference for a disjoint interpretation in response to both (60) and (61) equally.

⁹ As with Experiments 1 and 2, the disambiguation point has been shifted forward 300 ms to account for saccade time (Fernald et al., 2008).

3.3.2 Results

Figure 6 shows the mean proportion of looking to the disjoint interpretation event across the timecourse of the trial.

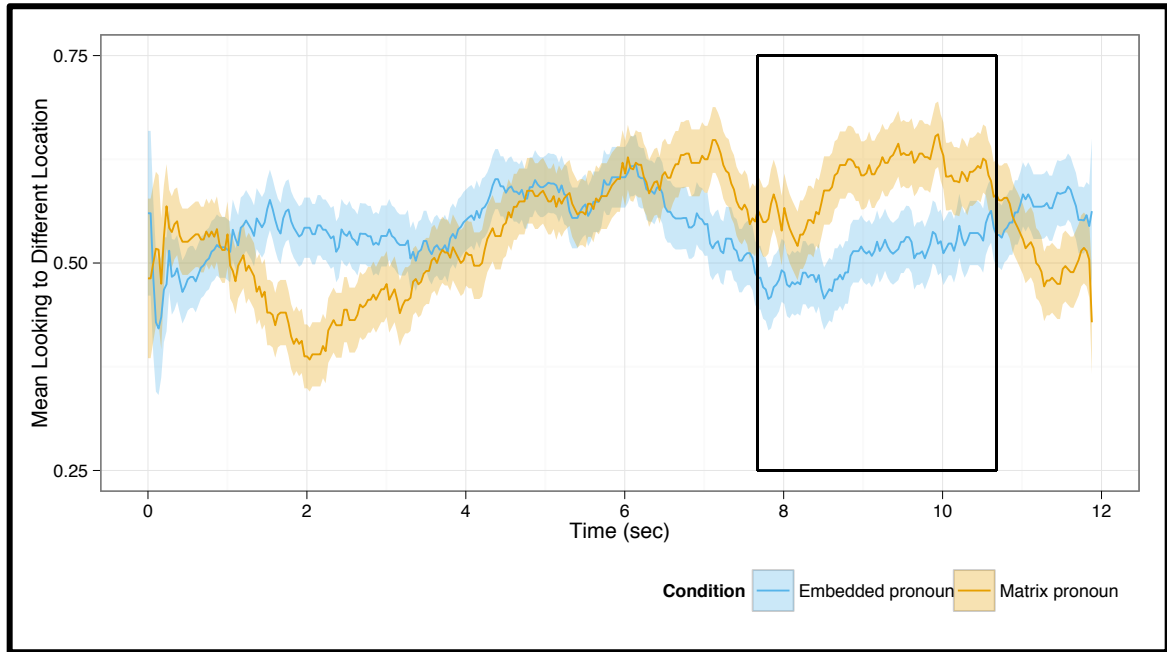


Figure 6 Looking behavior in Experiment 1. Outlined box represents critical windows for analysis

During the critical window following the onset of the second test sentence, children looked significantly more to the disjoint interpretation action in the MATRIX PRONOUN condition, while showing no preference in the EMBEDDED PRONOUN condition. This performance is reflected by a significant main effect of condition ($\beta=0.08165$, $SE=0.03965$, $p<0.05$). Planned comparisons within each condition show that looking to the disjoint interpretation event is significantly greater than chance in the MATRIX PRONOUN CONDITION ($t(25)=2.4$, $p<0.05$), but is not different from chance in the EMBEDDED PRONOUN CONDITION ($t(25)=0.5328$, $p=0.6045$).

3.3.3 Discussion

The results from Experiment 3 showed that at 30 months, children's interpretations of sentences like (39) and (40), repeated here as (62) and (63), show a consistent asymmetry. Children prefer to interpret (62) as referring to an event corresponding to a disjoint interpretation, while they have no preference for interpreting (63) as disjoint or coreferential. This behavioral pattern cannot be accounted for with only a bias against all cases of backward anaphora.

(62) She's painting the house that's on Katie's lap.

(63) The house that she's painting is in Katie's lap.

One additional aspect of performance is worth mentioning here. Notice that in the first 3-4 seconds of the trial, there is a sizable asymmetry in looking between conditions: children in the MATRIX PRONOUN condition show an initial preference for the coreferential interpretation, while those in the EMBEDDED PRONOUN condition show no early preference. Given that this is the exact opposite performance than what we predict for the critical window, this may at first glance be concerning. However, given our efforts to bias toward a coreferential interpretation and findings from adult sentence comprehension, this may not be entirely surprising. Adults have been shown to prefer interpreting a pronoun in the subject position as coreferential with the subject of the previous sentence (Arnold et al., 2000; Crawley & Stevenson, 1990; Gordon et al., 1993; Stevenson, Nelson & Stenning, 1995). Recall that the subject of the previous sentence in this case would be the character who is mentioned in the test sentence (corresponding to a coreferential

interpretation). Given that the subject of the test sentence is a pronoun in the matrix pronoun condition but not in the embedded pronoun condition, the preference to initially interpret this subject pronoun as referring to the most recently mentioned character and respond in looking to the coreferential interpretation event is in fact expected. One might imagine that the fact that this initial interpretation is later overridden, as shown in the critical window of analysis, makes the contrast between performance for these two sentence types all the more apparent.

These results replicate the findings of Experiment 1 and extend those of Lukyanenko, Conroy & Lidz (2014), showing children to exhibit ‘Principle C effects’- preferences for disjoint interpretation in Principle C contexts- at 30 months. Further, the results of Experiment 3 extend these findings to show that children do *not* exhibit such behavior to all backward anaphoric structures. This suggests that children’s performance is more consistent with knowledge of Principle C than it is with an interpretive bias based in the linear order of nominal elements.

3.4 General Discussion

Experiment 3 further supports the conclusion that restrictions on interpretations in Principle C contexts are in place by 30 months, and additionally presents new evidence that interpretations are not uniform across all backward anaphoric contexts. As such, these restrictions are more consistent with accurate knowledge of Principle C than with an interpretive bias based in linear order of nominal elements. However, two possible challenges may be raised to offer alternative accounts for such results; I examine each of these challenges in turn here.

One challenge in interpreting the results of Experiment 3 is that an additional asymmetry between the sentence types exists beyond the c-command relations between the pronoun and the R-expression. There is also an asymmetry in the type of relative clause exhibited in each sentence type. Subject gap relative clauses are those where the *filler*, the head of the relative, is displaced from the subject position of the embedded clause, as in (64); the ‘gap site’ from which the filler is displaced is identified by ‘___’. Object gap relative clauses, similarly, are those where the filler is displaced from the object position of the embedded clause, as in (65).

(64) I saw *the boy* that ___ kissed the girl.

(65) I saw *the boy* that the girl kissed ___.

Research on parsing of filler gap dependencies such as those exhibited in relative clauses has shown that even adults exhibit a significant slow-down in interpretation of object gap structures in comparison to subject gap structures (Gibson, 1998; Gordon, Hendrick & Johnson, 2001). When this finding is applied to the asymmetry in responses to sentences like (39) and (40), repeated here with the gap sites included as (66) and (67), an alternative explanation for the response pattern is evident.

(66) She’s painting the house that ___ ‘s in Katie’s lap.

(67) The house that she’s painting ___ is in Katie’s lap.

Chance performance in the case of (32), which exhibits an object gap relative clause, could be accounted for by a delay (or complete absence) of successful interpretation of the relative clause, rather than a successful interpretation in an unconstrained context. However, Kidd et al. (2007) show that 3-4 year old children seem to find object relatives with a pronoun in the subject position of the relative just as easy as subject relatives. Roland et al. (2012) find similar effect in adults, showing that given the appropriate discourse context, object relatives are no more difficult to parse than subject relatives.

While children have been shown to be capable of appropriately computing interpretations for relative clauses by 30 months (Lidz & Gagliardi, 2010), this competence is fragile (Gagliardi, Mease & Lidz, submitted) and deployment of this underlying competence could potentially be interrupted by additional processing demands, such as the reference resolution requirements present in (66) and (67). Future research will therefore explore alternative contexts in which c-command can be manipulated across backward anaphoric structures.

3.5 Conclusion

Children have been shown to utilize linear order of words to interpret sentences at early stages of language development. This suggests that children could in principle hypothesize a linear constraint on anaphora as a means to approximate Principle C. Additionally, backward anaphoric structures have been shown to be held to strict discourse conditions. This suggests that children could potentially have an initial overall preference for disjoint interpretation in backward anaphoric contexts that masks Principle

C effects. Either of these interpretations of a possible Linear Bias predicts identical performance to adult-like knowledge of Principle C in the case of simple single clause structures like those tested by Lukyanenko, Conroy & Lidz (2014).

In this chapter, I sought to further address the question of what interpretive mechanism drives children's early interpretations in Principle C contexts. I explored 30 month-olds' interpretations of backward anaphoric sentences which are unconstrained by Principle C, making both a coreferential and disjoint interpretation available. The results of Experiment 3 demonstrated that when Principle C blocks a coreferential interpretation, children exhibit a preference for a disjoint interpretation, but in backward anaphoric contexts where Principle C does not restrict interpretation, children show no preference for either interpretation. A bias to interpret all cases of backward anaphoric structures as disjoint in reference is not able to account for children's asymmetric responses across these sentence types. This experiment is the first exploration of 30 month-olds' interpretations of backward anaphoric structures that are unconstrained by Principle C, and the results suggest that interpretation even at this young age is more consistent with knowledge of Principle C than a simple bias based in linear order.

4 Identifying Knowledge Through the Contribution of Processing

4.1 Outline

Chapter 4 examines a novel analytic approach to identifying characteristics of knowledge underlying behavior. I utilize independent measurements of processes that may be implicated in deployment of knowledge in order to determine the nature of the knowledge that is being deployed. I exploit the predicted dependency between children's speed of processing syntactic information and their speed of interpreting sentences in Principle C contexts. In three preferential looking tasks, I use measures of processing at the lexical and syntactic levels to compare to performance on a Principle C task. I show that individual variation in the speed of interpretation in Principle C contexts is predicted by individual variation in syntactic processing speed. This finding suggests that the mechanism responsible for interpretation in Principle C contexts is dependent on syntactic composition.

4.2 Introduction

So far, we've discussed explorations of early Principle C effects which seek to identify the underlying knowledge by comparing differing predictions about behavior made by possible interpretive strategies. In Chapter 2, we explored a linguistic context in which Principle C and a bias for interpretation in transitive clauses predicted differing performance, and showed that children's behavior is more consistent with knowledge of Principle C than with a transitivity bias. In Chapter 3, we explored a context in which

Principle C and a linear order interpretive strategy predict differing performance, and showed that children's behavior is more consistent with knowledge of Principle C than with a linear interpretive strategy. While this method of comparison is effective for ruling out individual mechanisms that have been proposed as alternatives to knowledge of Principle C, it fails to directly implicate knowledge of Principle C. In other words, I have so far been able to what 30 month-olds' interpretive strategy is *not*; nothing in the findings thus far directly identifies what the interpretive mechanism *is*. In this chapter, I take a different approach in order to more closely identify specific characteristics of the underlying knowledge that drives behavior. Specifically, I explore the inherent dependency between grammatical knowledge and the deployment processes required to implement this knowledge; recognition of this dependency allow us to generate predictions about variance in behavior based on variance in implementation. We can draw conclusions about children's grammar by acknowledging that differing knowledge states will require different mechanisms for implementing that knowledge, and by finding ways of explicitly measuring the online implementation of such knowledge.

This research also serves a second goal, which is to understand more about the developmental trajectory of Principle C knowledge in children by identifying what factors beyond simply knowledge of the constraint contribute to interpretation in Principle C contexts. Even if Principle C is innately specified, for it to be implemented in the interpretation of a particular sentence children have to have learned enough about their native language to be able to identify which structures it will apply over. Understanding which factors affect early Principle C effects is important to understanding what aspects of grammatical competence are implicated in interpretation.

The research presented here explores the development of children's interpretations of sentences in Principle C contexts, demonstrating that individual variation in performance is predicted by variation in speed of processing at the syntactic, but not the lexical level. This relation between structural processing and interpretation in Principle C contexts suggests that children's comprehension is dependent on parsing of hierarchical structure. Because this correlation is not predicted by non-structural interpretive mechanisms, these findings suggest that children's knowledge of Principle C is in place by as young as 30 months of age. This work serves to demonstrate that investigation of individual differences at the processing level can be used as a probe into similarities across children at the knowledge level.

Measuring interpretation at the individual level

In this chapter, I explore individual differences in interpretation as a means of identifying the mechanism driving this interpretation. Specifically, I take advantage of the fact that different interpretive strategies will require different processes in order for interpretation to occur. Depending on how the restriction on interpretation is defined in children's grammar, different components will be necessary for applying that restriction to a given sentence. For example, if Principle C is the mechanism that drives the restriction on interpretation, the ability to compute c-command relations between an R-expression and c-commanding nominal expressions is necessary to be able to apply Principle C knowledge to interpreting a sentence. If, however, the mechanism driving interpretation were instead a linear order strategy, computing c-command relations would not be relevant to interpretation.

The question that arises is how we might identify differences in interpretation at the individual level, in order to identify whether differences in the interpretive process are explained by differences in the various implementation factors. The primary measure of interpretation in Preferential Looking tasks is defined by the proportion of looking time spent attending to a target image, which corresponds to a particular interpretation of a sentence. While proportional measures are useful for looking at the data in aggregate, the conclusions that can be drawn from the results are in some sense limited. Whether children reliably attend more to one image than another across a span of time can be taken as an indication that they have or lack some ability to comprehend a linguistic stimulus. However, this type of measure says nothing about which components of linguistic processing are implicated in allowing children to arrive at the interpretation they do. Further, this type of measure is unable to distinguish the type of knowledge driving this behavior; rather it can only show that children arrive at a certain type of interpretation, without directly identifying the knowledge that drives that interpretation.

A more effective measurement of individual differences in interpretation for our purposes is the speed with which an individual child arrives at a particular interpretation. Because we can identify the point at which children have all the information they need from the input sentence to begin to process the restriction on interpretation, we can identify how long after that point interpretation occurs. In essence, we are deriving a reaction time measurement that identifies the speed with which interpretation of a sentence is restricted to one of two possible meanings.

In order to derive a reaction time measure from Preferential Looking Paradigm data, we turn to research in the field of word recognition and word learning. Fine-grained

temporal analysis of PLP data has largely been developed by Anne Fernald and colleagues, in their research probing young children's developing abilities in online word recognition. This research does not measure children's syntactic knowledge, but it has provided a system of analysis of the fine-grained temporal structure of children's responses.

Fernald et al. (1998) first explored the timing implications of children's online comprehension of linguistic information by probing the change in recognition of object names across the second year of life. Children were presented with two familiar objects and a sentence that asked children to find one of the two objects (e.g. *where's the ball?*). Fernald et al. found that while all children were able to find the target item, the efficiency with which children did so varied by age, with the oldest children being able to attend to the target object even before word offset.

Swingley, Pinto & Fernald (1999) extended this result by focusing on word processing at 24 months and utilizing a dependent measure that takes individual shifts in attention into account. The task design was the same as Fernald et al. (1998), except that in one condition, the object names had an initial phonological overlap (e.g. *doggie/doll*), while in the other they did not (e.g. *doggie/tree*). Swingley et al. measured children's reaction time by focusing on trials where the child was attending to the distractor image at the onset of the target word; from these trials it is possible to measure the exact time it takes for the child to revert attention to the target image (see Fernald et al. 2008 for a detailed explanation PLP tasks designed to utilize online measures). Results showed that response latency was significantly longer in the overlap condition, suggesting that even very young children process speech incrementally, and adjust their attention accordingly.

I extend these results here to show that this incremental processing can be scaled up to the level of incremental syntactic composition and interpretation. I utilize the latency measure identified by Swingley et al. to compare speed of interpretation in Principle C contexts at the individual level.

Factors affecting interpretation

This timecourse analysis allows us to capture the speed of interpretation in Principle C contexts, which we can use to infer the underlying computations required to form such interpretations. We have noted already that due to the indirect relation of observed behavior to the knowledge that motivates that behavior, it is challenging to reason directly from performance about the nature of children's grammatical knowledge. However, it is possible to take advantage of the mediating effect of the processes required to deploy grammatical knowledge, by making predictions about how different knowledge states will be deployed. Depending on children's developing syntactic knowledge, they may arrive at an interpretation for a sentence in a number of different ways. As adults, the primary mechanism for such interpretation will be syntactic composition, however it's possible that children may be able to form an interpretation via non-structural methods (such as mechanisms similar to those discussed in Chapters 2 and 3). If children have adult-like grammatical knowledge of Principle C, their interpretation in Principle C contexts is restricted by recognition of the c-command relation between an R-expression and another nominal element. If c-command holds, any interpretation where an NP is co-indexed with the relevant R-expression is illicit. However, if children derive a preference for disjoint interpretations in such contexts via a non-Principle C mechanism, then the

deployment of their final interpretation will not be inherently tied to syntactic processing as it would be with the structural computation inherent in c-command relations.

Because each type of interpretive mechanism will require different processes to deploy interpretation, we can make predictions about the relationship between children's behavior and identifiable processing mechanisms. Specifically, we predict children's behavior to be subject to their syntactic processing abilities only if they derive an interpretation by syntactic means. This dependency, schematized in Table 1, allows us to form concrete predictions about children's performance. We expect that children's performance in Principle C contexts will be predicted by their overall abilities in processing hierarchical structure, but only if the knowledge driving interpretation in Principle C contexts is syntactic in nature.

Knowledge	Principle C constraint	Alternative interpretive constraints
Deployment	Dependent on ability to process hierarchical structure	Independent of ability to process hierarchical structure
Behavior	Variation predicted by variation in syntactic processing speed	Variation independent of variation in syntactic processing speed

Table 1 Predicted dependencies between deployment processes and observed behavior by hypothesized interpretive mechanisms.

In this chapter, I explore several possibilities for factors relevant to implementing interpretations. First, as it has been identified to affect children's performance in Principle C contexts (at least in some results), is vocabulary size. One of Lukyanenko et al.'s chief findings was that the size of children's productive vocabulary predicted performance in Principle C and reflexive contexts. The results presented in Chapters 2 and 3 found no overall effect of vocabulary size on performance, however it could be the

case that methodological differences between these tasks allowed the effect of vocabulary to be shown in the results of LCL but masked the effect in our follow-up studies. Therefore we include the measure of MCDI vocabulary score as one potential factor that could affect children's interpretation in Principle C contexts.

It is as of yet unclear the exact role of vocabulary as a predictor of performance. No account predicts that vocabulary itself should directly affect the acquisition of Principle C¹⁰; it is unclear how the size of a child's lexicon would bear any direct relation to constraints on anaphora. For this reason, it seems that vocabulary may be the surface index of a different underlying mechanism (or mechanisms), for which variability more straightforwardly predicts variability in performance with Principle C. LCL suggest two possibilities. One is that vocabulary could reflect some aspect of children's grammatical development, such that the absence of Principle C effects in the low vocabulary children stems from their lacking some aspect of grammatical competence which allows successful application of the constraint. For example, if children with smaller vocabularies do not yet command the structure of transitive clauses, then they will be unable to successfully build a structure for the sentences and compute c-command relations to determine if binding conditions hold. The possibility that the vocabulary effect stems from different levels of grammatical development is supported by research showing vocabulary to be indicative of children's grammatical development (Marchman & Bates, 1994; Dale et al. 2000; Devescovi et al., 2005). A second possibility is that

¹⁰ One exception would be that low vocabulary children could be predicted to fail on such a task if they do not know the verbs that were used in Lukyanenko et al.'s sentences; however this seems unlikely, both because the verbs used were highly common actions (cover, dry, fan, paint, pat, spin, squeeze, wash), and each action was introduced separately prior to the test phase.

vocabulary is an index of children's speed of processing, such that the failure of low vocabulary children to show Principle C effects in LCL's results is a result of their inability to complete the interpretation and mapping processes quickly enough for this task. This possibility is supported by research in the word recognition domain, showing that vocabulary is related to children's speed of processing (Fernald, Perfors & Marchman, 2006 and Hurtado, Marchman & Fernald, 2008, among others)¹¹.

The second factor that I consider as possibly having an effect on interpretation is that of processing speed. The concept in itself is vague; I will first consider the term as it is utilized in the word recognition domain, as defined in Swingle et al (1999). In the context of a task which presents two common objects and a sentence which asks children to find one of the objects, a measure of response latency can be seen as measuring the speed of interpreting the target noun and mapping that interpretation onto the corresponding visual scene. In effect, this measure can be understood as a measure of children's early speed of lexical access. I will this refer to this measure in the remainder of the chapter as Lexical Access Speed (LAS). While lexical access will inherently be required for any method of interpreting sentences in Principle C contexts, individual variation with respect to LAS is a factor that is more likely to explain variance in speed of interpretation only if it is one of the primary components contributing to interpretation. That is, we expect processing of syntactic information (such as the relevant c-command relations) to be most critical to interpretation with knowledge of Principle C; however,

¹¹ However, these findings are not entirely uncontroversial; several studies have found little or no evidence for a significant relation between vocabulary and processing speed. Appendix D summarizes studies that report the significance of vocabulary size on response latency. Note that even within studies, results can vary by age.

processing of lexical information could be most critical to interpretation with an alternative non-structural mechanism.

Because we expect processing at the syntactic level to be inherently important for interpretation under a Principle C account of performance, I also consider syntactic processing speed as a potential factor affecting interpretation. While the processing speed measure derived from the LAS task is widely utilized in word recognition research, a comparable standard measure for syntactic processing speed does not yet exist; I have therefore designed a Phrase Structure Integration Speed (PSIS) task in order to generate such a measure. The goal in creating a measure of children's PSIS was to present a linguistic context representable in PLP format which would require children to compute the hierarchical structure of a phrase in order to be able to identify the intended meaning, rather than being about to rely on lexical information alone. To accomplish this, I utilize superlative constructions as in (68), with a corresponding visual array as in Figure 7.

(68) Where's the biggest red train?

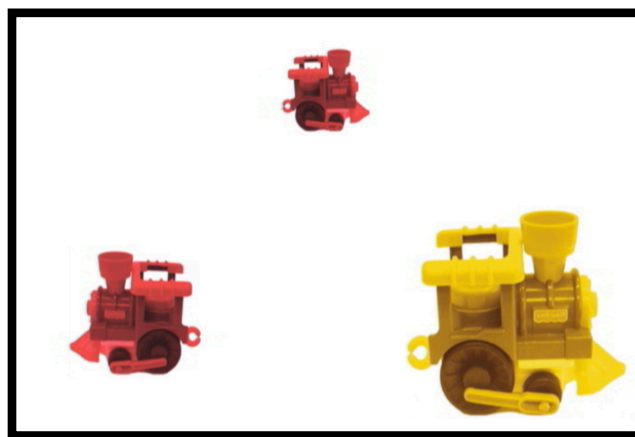
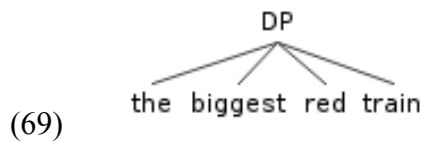
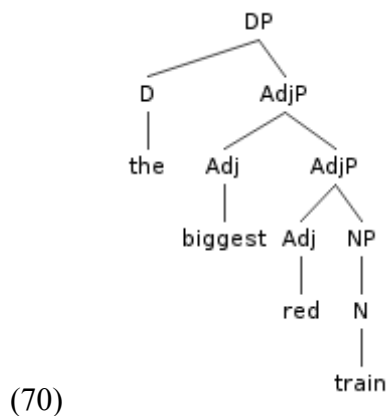


Figure 7 Sample visual array for Phrase Structure Integration Speed task. Largest item is yellow; smaller two items are red.

Given the visual array presented in Figure 7, it is clear that there is no one item that could be identified as the target by simply identifying the relevant features in absence of a structured representation. Let us consider how interpretation would occur if no internal structure was applied to the phrase, as in (69).



If children do not attribute hierarchical structure to the phrase, then each element would be interpreted conjunctively, and the target item would be identified as one which satisfies the combination *biggest + red + train*. None of the three items pictured satisfies all of the features *biggest + red + train*, because the item that is globally biggest does not satisfy the feature *red*, and the items that are red cannot be interpreted as satisfying the feature *biggest*. In order to arrive at an adult-like interpretation of (68), children would need to represent the NP *biggest red train* hierarchically, as in (70).



Given this hierarchical structure, the phrase [red train] can be interpreted as a unit to which the superlative *biggest* applies. With this interpretation, the biggest item in the set satisfied by the features *red* + *train* is the target. Thus with this superlative construction, we are able to identify a simple case in which accurate interpretation can be attributable to the use of hierarchical structure. In this way, the speed of interpretation on this task can be considered a measure of processing syntactic information.

The current study

As identified above, the goal of this study was to compare individual variation in various independent measures of linguistic ability to the individual variation in speed of interpretation in Principle C contexts. We predict that if the knowledge driving behavior is Principle C, then speed of interpretation should be predicted by speed of processing syntactic information.

The research presented here is composed of three tasks that allow us to test this prediction. A Principle C task like that of Lukyanenko, Conroy & Lidz (2014) determines children's interpretations in Principle C contexts. The remaining two tasks have been designed to measure children's processing speed, in order to compare this to children's performance on the Principle C task. The first is a task modeled off those used in word recognition research, which generate measures of lexical processing speed. The second is a similar task designed to generate measures of children's syntactic processing speed.

4.3 Experiment 4

4.3.1 Method

4.3.1.1 Participants

We tested 64 English-speaking children (32 males) 28-32 months of age (range = 28;2-31;28; median = 30;7; mean = 30;6) recruited through the University of Maryland Infant and Child Studies Database. Six additional children were tested but were excluded from the final sample for the following reasons: failure to complete all three tasks (n=2); equipment failure/ experimenter error (n = 4).

MacArthur-Bates Communicative Development Inventory (MCDI) Words and Sentences long forms were collected for each child, revealing a range of vocabulary sizes from 99 to 680 words (median = 562; mean = 514). With the goal of comparing individual performance across tasks, each participant was tested on each of the three tasks described below. Tasks were completed in one session that lasted around 30 minutes (including play breaks between tasks when needed).

4.3.1.2 Materials & Design

Task 1: Principle C Task

The task designed to test interpretations in Principle C contexts used identical stimuli and design to that of Experiments 1 and 2, with the exception of audio stimuli. See Section 2.3.1 for a description of the task design and visual stimuli. Audio stimuli followed that used by Lukyanenko, Conroy & Lidz (2014). During the Familiarization phase, children heard audio which described the action, but was ambiguous as to the

agent and patient of the event, as in (71). This was to allow either the self-directed or other-directed events to be a possible representation of the test sentences.

(71) Wow! There's Anna and Katie! It looks like somebody's getting patted!

Oh look! There they are again! Somebody's getting patted again!

During the Test phase, children heard three repetitions of the test sentence in different frames, as in (72) and (73) below.

(72) She's patting Katie! Do you see the one where she's patting Katie? Find the one where she's patting Katie!

(73) She's patting Katie! Do you see the one where she's patting Katie? Find the one where she's patting Katie!

Children were presented with a total of 8 test trials in a between subjects design: half of the subjects heard name condition sentences as in (72) for all trials, and half heard reflexive condition sentences as in (73). For each iteration of the test sentence, the audio was aligned by the onset of the object NP for later analysis, as this is the critical point when an interpretation strategy based in Principle C could begin to identify the intended interpretation of the sentence. This resulted in a 2 (Condition) X 3 (Window) design.

Task 2: Lexical Access Speed Task

In addition to a task probing interpretations in Principle C contexts, we included two additional tasks designed to elicit measures of participants' lexical and syntactic processing speed. The Lexical Access Speed (LAS) task, which generated measures of each child's lexical processing speed, was a word-object mapping task modeled after that of Swingley, Pinto & Fernald (1999). Children were presented with two images of common objects, and a sentence which then directed children to find one of the two objects. Objects presented were chosen from the most common nouns in young children's vocabularies (all words, listed in Appendix C, are reported to be said by at least 90% of 30 month old children). Figure 8 presents a sample array. After observing the images in silence for approximately 1 second, children heard two instances of the test sentence, naming one of the two items, as in (74).

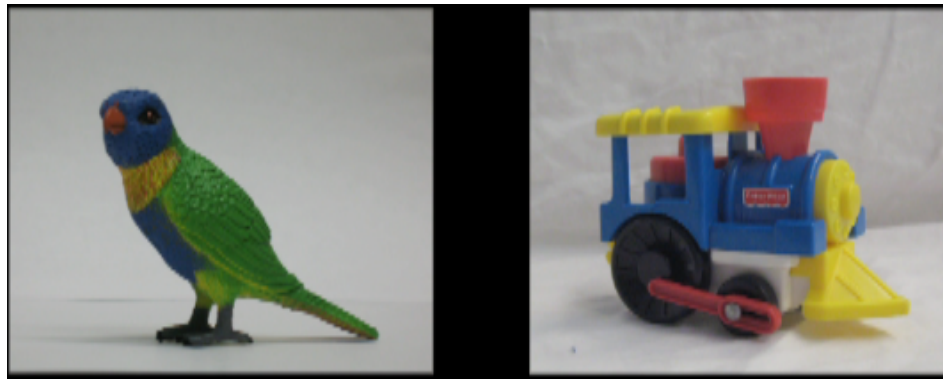


Figure 8 Experiment 4 Lexical Access Speed Task sample array

(74) Where's the train? See the train?

Each of the 8 trials lasted a total of 5 seconds. Position of the target object was counterbalanced across trials; target object and order of presentation (2 possible lists) were counterbalanced across subjects.

Reaction time (RT) was calculated by determining the latency to attend to the target image on distractor-initial trials (i.e. those trials where the child was attending to the distractor image at the onset of the target word). These RT values were then averaged across trials to derive a LAS measure for each participant.

Task 3: Phrase Structure Integration Speed Task

While the processing speed measure derived from the LAS task is widely utilized in word recognition research, a comparable standard measure for syntactic processing speed does not yet exist; we have therefore designed a Phrase Structure Integration Speed (PSIS) task in order to generate such a measure. We maintained the word-object mapping task design in order to keep task demands as comparable as possible to the LAS task. The visual array included three objects instead of two, which required a complex NP as the description of the target object. As discussed in Section 4.2 above, children must treat the NP as hierarchically structured to accurately interpret the description and locate the correct object; in this way, this new task allows for a measure of the speed of processing syntactic structure. Children were presented with three images of all the same kind (e.g. three trains); objects were again drawn from the most commonly known nouns (see Appendix C). The three objects in each set varied in size and color, with the two smaller items being colored the same. See Figure 7 above for a sample array. Each of 24 arrays contained a different set of objects. A constant size ratio of 3 : 4.5 : 7.5 was maintained

between the smallest, medium, and largest item. This ratio was chosen so that the smallest item was not so small as to not be easily identifiable, so that the largest item was contained within its quadrant of the screen (to facilitate accurate coding), and so that the medium item was differentiable from the smallest item, yet significantly smaller than the largest item so that it could not be considered big by itself. After observing each array in silence for approximately 1 second, children heard an introductory sentence which identified the type of objects in the array (e.g. *Oh look! Now there are some trains!*).¹² Children were then presented with the test sentence. In 12 of 24 trials, the sentence contained the superlative *biggest* but no color adjective, as in (75), indicating the globally largest item. In 12 trials, the sentence contained a color adjective, as in (76), indicating the larger item in the subset of the two similarly colored items. SUPERLATIVE trials were 8.37 seconds long; SUPERLATIVE + ADJECTIVE trials were 8.77 seconds long. Order of item presentation was counterbalanced across subjects. Position of the target item, color of the target and distractor items, and sentence type (SUPERLATIVE or SUPERLATIVE + ADJECTIVE) were counterbalanced across trials. Additionally, each of these factors was varied pseudo-randomly and interspersed throughout the task relatively evenly, to avoid results driven by task effects¹³. Reaction times were calculated in the same way as for the LAS task to derive PSIS measures for each participant.

¹² Fernald, Thorpe & Marchman (2010) showed high rates of ‘false alarming’ shifts to the distractor image when 30 month-old children were presented with two objects of the same kind and an adjective + noun phrase picking out one of the items (e.g. *where’s the blue car* in the context of a blue and a red car). Therefore we included this introductory sentence to help ameliorate processing of the target noun.

¹³ The target item never occurred in the same position on consecutive trials. The same sentence type occurred in no more than two consecutive trials. The same color was

(75) Where's the biggest train?

(76) Where's the biggest red train?

A size adjective was used because it allowed the strict controlling of the ratio between the three objects in the array, and across trials, which is less feasible with other adjectives (e.g. specifying levels of fuzziness for *the fuzziest blue cat*). The particular size adjective *biggest* was used because *big* has been found to be the most frequently used base form for both comparatives and superlatives in young children's production (Layton & Stick, 1979). This coincides with data from the MCDI Lexical Norms Database showing *big* to be one of the earliest prominent adjectives in young children's expressive and receptive lexicons (Dale & Fenson, 1996). We avoided using *smallest* due to research showing that children acquire the positive dimensional adjectives before the corresponding negative adjective (Donaldson & Balfour, 1968; Ehri, 1977). The color words *red*, *blue*, *yellow*, and *green* were used due to their being the most common color words in children's productive vocabularies by this age (with the exception of *orange*, which was not used because of the potential confounding of the color name with the noun).

In considering individual differences in processing at the syntactic level, it is important to recognize that children may differ in two aspects required for processing and interpreting structural information. First, they may differ simply in how quickly they are able to integrate new lexical information into the structure they have begun to build

presented in no more than three consecutive trials, or presented in the same position in no more than two consecutive trials.

(Fernald, Thorpe & Marchman, 2010). This integration aspect can be thought of as the level of incrementality with which children parse a structure; children who are faster at incorporating new lexical items into their structure will be those who are more efficient in either generating new structural elements or slotting lexical information into predicted structure (or both).

An additional aspect critical to syntactic processing is children's ability to deal with new lexical information that doesn't fit into their predicted structure. If children incrementally incorporate information into a structure and predict the most likely completion of such a structure, they may boggle when confronted with new lexical information that does not fit into the structure they have predicted (Trueswell et al., 1999; Snedeker & Trueswell 2004; Lidz, White & Baier, submitted). At this point, children will be required to alter their predicted structure to accommodate the new information. So in addition to differing simply in their ability to generate structure and incorporate information into a structure, children may also differ in their ability to accommodate information that is in conflict with their initial interpretation. Crucially, this additional revision aspect may affect children differentially, depending on their overall abilities at building and interpreting structure: children who are slower overall to incorporate information into a structure may be less efficient at predicting likely completions of the structure, and therefore may be less likely to even need to revise such predictions. Contrastively, children who are quicker to integrate new lexical information may be those who use a more predictive strategy, and may be more prone to have to revise an initial interpretation.

Given that speed of processing syntactic information comprises two distinct capabilities, our PSIS task allows us to measure each of them. Consider first the SUPERLATIVE condition sentences like (75). Assuming an incremental parsing strategy, children hear the superlative *biggest*, access the lexical entry, and begin to build a structure, even without all the information they will need to complete it. They will then assign an interpretation over this incomplete structure, predicting it to be referring to the biggest item in the set, which can be mapped to the visual array to pick out the biggest item. When they hear *train*, they will update their interpretation to finding the biggest item *in the set of trains*. This picks out the same item in the array, so their original interpretation and mapping will be confirmed as correct.

Alternatively, if children are less incremental in parsing, they may wait until they've built the full structure before interpretation. This will result in the same behavior as more incremental parsers but with a possibly different timecourse. Thus differences in speed may reflect differences in the incrementality with which children perform the interpretive process.

In the SUPERLATIVE + ADJECTIVE condition, as in (76), these two parsing strategies are more distinct in the behavior they predict. In these trials, incremental parsers will again assign an initial interpretation that picks out the biggest item upon hearing the word *biggest*. However, because the following word is a color adjective, such as *red*, children who are parsing incrementally will have to *revise* the initial interpretation to one which picks out the largest item *from the subset of red items*, which will map to the larger of the two red items. Thus there is a crucial difference between the processes required in these

two conditions of the PSIS task, in that the SUPERLATIVE + ADJECTIVE condition potentially requires this revision of the initial interpretation.

4.3.1.3 Procedure

Procedure for running subjects was identical to that of Experiments 1-3, with one exception. In coding the data from the PSIS task, data was coded to indicate whether children were attending to the left, right, top center, or not at all. Inter-coder reliability remained high; across three coders, percent agreement was above 97% in all cases, with Cohen's kappa scores of .95 and above.

4.3.1.4 Analyses & Predictions

The analyses presented below are designed to address a series of questions about children's performance, outlined below. Given the range of effects we wish to explore, these questions are answered through several different types of analyses, described in the following sections.

(Q1) Do children successfully distinguish interpretations in Principle C and reflexive contexts?

Addressed by Timecourse Analyses

(Q2) To what can the vocabulary effect observed by LCL be attributed: grammatical knowledge, processing speed, or experimental design?

Addressed by Timecourse Analyses and Processing Speed Analyses

(Q3) Do children successfully interpret hierarchical structure in the PSIS task?

Addressed by Timecourse Analyses

(Q4) Does our new measure of syntactic processing speed capture something distinct from the (lexical) processing speed measure formalized by Swingle, Pinto & Fernald (1999)?

Addressed by Processing Speed Analyses

(Q5) Which of these factors, if any, predict performance in Principle C contexts? What inferences can we therefore make about the mechanism driving Principle C effects?

Addressed by Growth Curve Analyses

Timecourse Analyses

The majority of the analyses presented here compare looking behavior across the timecourse of the test trial in each task. Performance is measured over the first 2 seconds following the disambiguation point. In the Principle C task, where children hear three repetitions of the test sentence, we form three corresponding windows of analysis comprising 2 seconds after the disambiguation point in each iteration of the sentence.

Analysis of overall performance on the Principle C task will address the main concern of children's interpretations in each condition. Before the point of disambiguation, we expect that in both conditions children should attend to the reflexive and non-reflexive events roughly evenly; by that point they will not have received any information about the interpretation of the test sentence. Following the disambiguation, we predict performance to vary depending on the interpretation children assign to the test sentence. In the REFLEXIVE condition, if children have an adult-like understanding of reflexives, we predict decreased looking to the non-reflexive event (corresponding to

increased looking to the reflexive event), indicating a reflexive interpretation of the sentence. In the NAME condition, if children have knowledge of Principle C, we predict increased looking to the non-reflexive event, indicating a non-reflexive interpretation of the sentence.

Finally, a comparison of performance by vocabulary score will address more closely the vocabulary effect observed by Lukyanenko, Conroy & Lidz (2014), in order to determine whether vocabulary is indicative of having or lacking a particular component of grammatical knowledge, as suggested by LCL. If this is the case, we predict performance for low vocabulary children to be at chance in both conditions, with performance for high vocabulary children significantly different.

Processing Speed Analyses

The main purpose of the LAS and PSIS tasks was to gain a measure of individual children's LAS, the 'processing speed' measure utilized in the word-learning literature, and PSIS, a processing measure more closely tied to interpreting structural rather than lexical information, as measures with which to compare children's Principle C behavior and hence evaluate the mechanisms engaged in comprehending sentences exhibiting Principle C effects. Additionally, comparisons to these measures allow us to determine whether either of these mechanisms could account for the vocabulary effect behavior observed by LCL. The speed measure standardly used in word-learning literature corresponds to the mean latency after the onset of the disambiguating word to re-orient to the target image on distractor-initial trials. In a preferential looking task, behavior upon

hearing the target lexical item depends on which image the child happens to be attending to at that moment. If the child is already attending to the target image, then ‘correct’ performance consists of maintaining this fixation; alternatively, if the child is attending to the distractor image when the sentence is disambiguated, then ‘correct’ performance requires shifting fixation toward the target image. Because a shift in attention is only necessary (or appropriate) on distractor-initial trials, it is from this subset of the data that our measure is derived. The point of disambiguation for each sentence type is defined as the point at which children have enough information to form an accurate interpretation of the sentence. In the LAS task, this point is the onset of the target noun. We formed two measures of Phrase Structure Integration Speed (PSIS): one measure over the SUPERLATIVE trials, and one over the SUPERLATIVE + ADJECTIVE trials. In the PSIS task, for SUPERLATIVE condition sentences, like in the LAS task, the disambiguation occurs at the target noun. In the SUPERLATIVE + ADJECTIVE condition sentences, the disambiguation occurs at the color adjective¹⁴. A final value for each individual child was made for each of these three measures by taking the average latency on all distractor-initial trials.

To assess relations between these processing speed measures (as well as their relation to the vocabulary measure), we compared each participant’s vocabulary, LAS, PSIS SUPERLATIVE, and PSIS SUPERLATIVE + ADJECTIVE values and measured correlation coefficients between the values for each pair of measures. Previous research on individual differences has also sometimes treated measures such as vocabulary as a group distinction, rather than analyzing values as a continuous measure (akin to LCL’s comparison of performance by high vs. low vocabulary children). For this reason, we also

¹⁴ As in Experiments 1-3, the disambiguation point has been shifted forward 300 ms to account for saccade time (Fernald et al., 2008).

analyze values on each measure as separated into two groups, defined by the median value for that measure. We then compare distributions of participants by median split groups, to determine whether children consistently fall into above-median or below-median groups across multiple measures. The comparisons of each of these measures will allow us to identify whether they are independent effects, so that we can use each independently to probe performance with respect to Principle C.

Growth Curve Analyses

The primary objective of the research presented here is to examine the predicted dependency between individual variation in speed of syntactic computation and variation in performance with respect to Principle C. If one component of the mechanisms engaged for interpreting sentences in Principle C contexts involves structure building, then we expect our measures of structure building (PSIS measures) to be correlated to performance in the Principle C task. In order to more closely investigate individual differences in response latency, we employ statistical analysis of growth curve modeling (Mirman, Dixon & Magnuson, 2008).

Common practice in the analysis of timecourse data such as that gained from preferential looking tasks has generally been to collapse across time for a given window of analysis; this practice yields proportions of a looking time measure over which ANOVA or t-test can be performed (indeed, we employ such methods here in the majority of our general analysis). However, this practice effectively wipes out the fine granularity of time course data; as our primary concern is the precise timecourse of children's interpretation, such methods are not ideal. Growth curve analysis is a statistical

method specifically designed to assess change over time (originally employed to analyze data from longitudinal studies, although the same methods can be used on a much smaller timescale).

This growth curve modeling approach is based on the assumption that observed fixation proportions reflect an underlying probability distribution of fixations. Two hierarchically related sub-models are employed to capture the observed fixation pattern. The level-1 sub-model captures the effect of time, and the level-2 sub-model captures the effect of individual variation. (77) presents an equation of the form $y = ax + b$ representative of a level-1 model¹⁵. This gives a value for the dependent measure Y (proportion of trials where looking was to the non-reflexive event) for a particular child i at time point j . In this model we have an intercept term α , a slope term β , and an error term ε . Both the intercept and slope terms vary across individuals, as captured in the corresponding level-2 model for each term, illustrated in (78a-b). In each of these equations, the first variable γ corresponds to the structural components (the value averaged across participants), while the second variable ζ corresponds to the stochastic components (the deviation of the individual's value from the average value). The subscripts on these terms correspond to the order of polynomial term to which the variable relates (0 referring to components of the intercept term and 1 referring to

¹⁵ An equation such as this would capture a linear increase in looking; more variation in the form of looking pattern (in other words, the shape of the fixation curve) can be captured with the inclusion of higher order polynomial terms; however, it is likely that this level of model will be appropriate for our analysis. Recall that we aim to capture the behavior corresponding to children's first shift in attention to the target event after the onset of the disambiguating word in the sentence. Correspondingly, a single shift in fixation will likely be best captured by a first-order (linear) polynomial.

components of the slope term). Thus the corresponding level-2 model can be defined as in (79).

$$(77) \quad Y_{ij} = \alpha_{ij} + \beta_{ij} * \text{time}_{ij} + \varepsilon_{ij}$$

$$(78) \quad \text{a. } \alpha_{0i} = \gamma_{00} + \zeta_{0i}$$

$$\text{b. } \beta_{1i} = \gamma_{10} + \zeta_{1i}$$

$$(79) \quad Y_{ij} = (\gamma_{00} + \zeta_{0i}) + (\gamma_{10} + \zeta_{1i}) * \text{time}_{ij} + \varepsilon_{ij}$$

This multi-level modeling approach allows us to capture both effects across time and across individuals; this analysis thus permits investigation of individual variation in response latency on the Principle C task, to determine if this variation relates to those individuals' values on our measures of syntactic processing speed, gathered in the PSIS task. The effect of these processing speed measures is evaluated by including them in the level-2 model, as shown below in (80). The terms γ_{0S} and γ_{1S} here index the effect of our values of processing speed (represented here as S) on the intercept and linear terms.

$$(80) \quad Y_{ij} = (\gamma_{00} + \gamma_{0S} * S + \zeta_{0i}) + (\gamma_{10} + \gamma_{1S} * S + \zeta_{1i}) * \text{time}_{ij} + \varepsilon_{ij}$$

In addition to comparing our measures of syntactic processing speed, we also considered our measure of LAS as well as measures of MCDI vocabulary. If children's behavior in Principle C contexts is driven by adult-like knowledge of Principle C and hence requires building a hierarchical syntactic representation, then deployment of their knowledge will depend on the speed with which children process syntactic information.

Similarly, adult-like interpretation with respect to sentences containing reflexives will also require structural computation in order to confirm a binding relation between the reflexive and its antecedent. If children rely on knowledge of phrase structure to interpret sentences, then the speed with which they interpret sentences will be dependent on the speed with which they deploy their structural knowledge. In this way, we predict individual differences in response latency in the Principle C task to be predicted by individual differences in syntactic processing speed, as measured by our two PSIS measures. Alternatively, if children rely on non-structural cues such as lexical information to interpret such sentences, then their speed of interpretation will instead be dependent on the speed with which they access lexical items. In this case, we predict response latency to be predicted by lexical processing speed, measured by the LAS measure. While a non-speed measure such as vocabulary size does not offer as straightforward a prediction, we may expect vocabulary to affect children's ability to access particular lexical items, especially those utilized in the test sentences. In this case, either the LAS measure or MCDI vocabulary could predict speed of interpretation.

4.3.2 Results & Discussion

Principle C Task

Questions Addressed

- (Q1) Do children successfully distinguish interpretations in Principle C and reflexive contexts?

(Q2) To what can the vocabulary effect observed by Lukyanenko, Conroy & Lidz (2014) be attributed: grammatical knowledge, processing speed, or experimental design?

Figure 9¹⁶ presents the results of the Principle C task for each condition. Before the point of disambiguation, behavior does not differ across conditions ($F(1,63)=0.6155$, $p>0.1$). Following disambiguation, we analyzed 2000 ms windows after the onset of the disambiguating word in each repetition of the test sentence (as indicated on the graph by the black boxes). A 2 (condition) x 3 (window) ANOVA reveals a main effect of condition ($F(1,63)=24.07$, $p<0.001$) a marginally significant main effect of window ($F(2,63)=2.329$, $p=0.0978$), and no interaction. Planned comparisons of behavior within each window show that looking to the non-reflexive image is significantly different from chance in the first and second windows for the NAME condition ($t(31)=4.475$, $p<0.001$; $t(31)=2.821$, $p<0.01$; $t(31)=0.9976$, $p>0.1$ for each window respectively), but only in the second window for the reflexive condition ($t(31)=-0.3637$, $p>0.5$; $t(31)=-2.411$, $p<0.05$; $t(31)=-1.388$, $p>0.1$ for each window respectively).

¹⁶ All graphs were generated using the ggplot2 library (Wickham, 2009) in R (R Development Core Team, 2010).

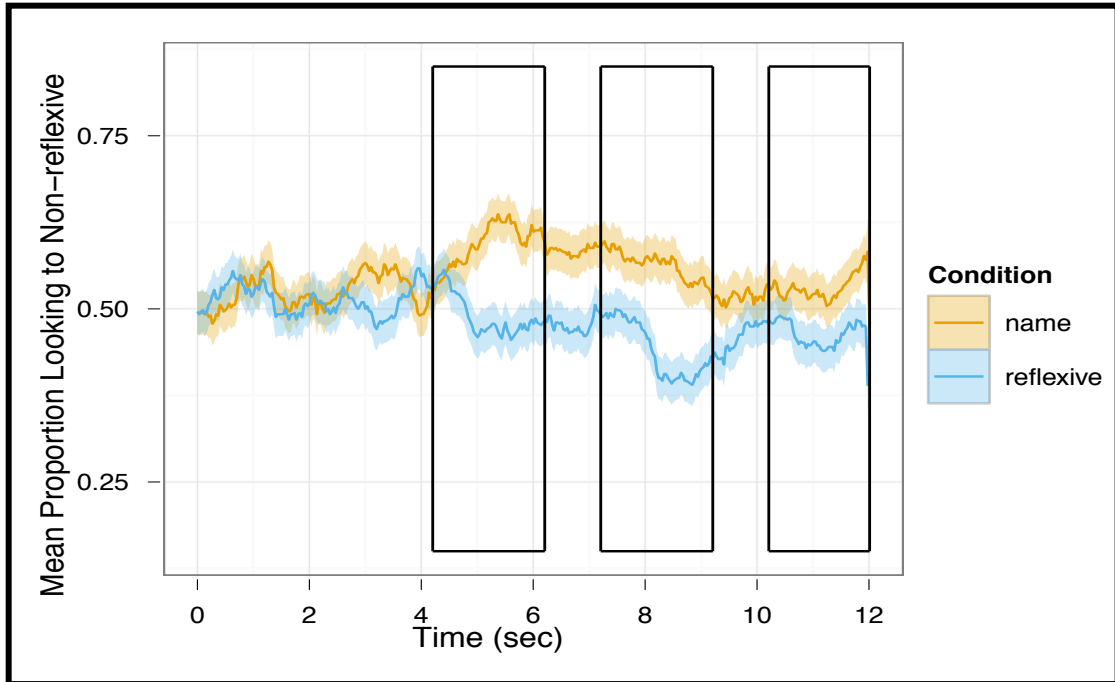


Figure 9 Looking behavior in Experiment 4 Principle C Task. Outlined boxes represent critical windows for analysis (2000ms following the onset of the disambiguating object NP in three iterations of the test sentence).

These results replicate the finding of LCL as well as those of Experiments 1 and 3 that by 30 months, children demonstrate a constraint on pronoun interpretation in Principle C contexts, looking preferentially to a disjoint rather than a coreferential interpretation.

Questions Answered:

(Q1) Do children successfully distinguish interpretations in Principle C and reflexive contexts?

(A1) **Yes**

To further explore the vocabulary effect observed by Lukyanenko et al., we analyzed performance by comparing behavior based on MCDI vocabulary. Figure 10 presents the performance in each condition of the Principle C task by MCDI vocabulary.

A 2 (condition) x 2 (vocabulary) x 3 (window) shows a significant effect of condition ($F(1,63)=24.12$, $p<0.001$) and a marginal effect of window ($F(2,63)=2.324$, $p=0.0982$), as well as a marginal effect of vocabulary ($F(1,63)=3.154$, $p=0.0759$). Additionally, we find an interaction between vocabulary and condition ($F(1,63)=3.928$, $p<0.05$). Comparisons within condition to explore the interaction term show a main effect of vocabulary in the REFLEXIVE condition ($F(1,63)=9.478$, $p<0.01$), and no such effect in the NAME condition ($F(1,63)=0.0174$, $p>0.5$).

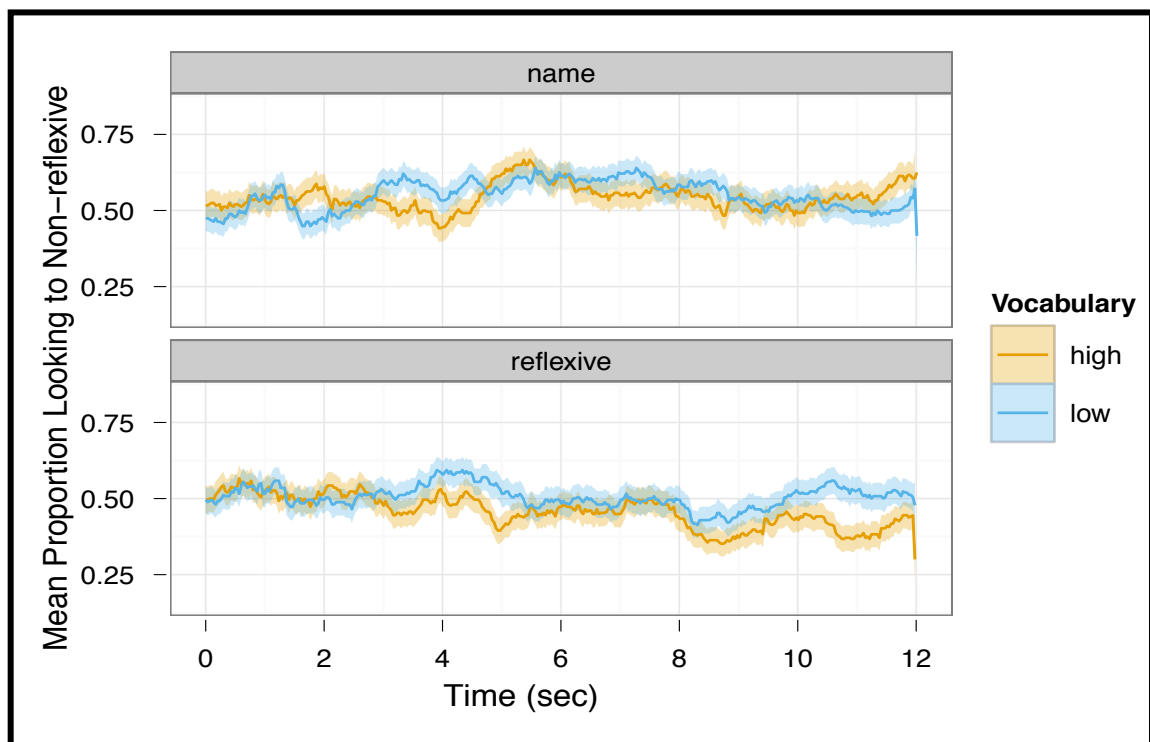


Figure 10 Looking behavior in Experiment 4 Principle C task by MCDI Vocabulary Score median split groups

The fact that we find an effect of vocabulary size in the REFLEXIVE condition but not the NAME condition is a partial replication of LCL's observed vocabulary effect. It also replicates the lack of a vocabulary effect in Principle C contexts as found in

Experiment 1. Overall, this result allows us to interpret LCL's finding in a new light. It seems that their vocabulary effect in the NAME condition could have been a spurious task-driven effect. As discussed in Chapter 2, this may represent a difference arising due to differing subject designs: while LCL used a within-subjects design, meaning that children heard four trials with each sentence type, we used a between-subjects design, where children heard eight trials all with one sentence type. Having to interpret multiple sentence types could have masked knowledge in the results of Lukyanenko et al., which becomes more pronounced when children are presented with only one sentence type across trials. However, the fact that the vocabulary effect maintains in the REFLEXIVE condition suggests that LCL's vocabulary result in the REFLEXIVE condition was not driven merely by a task effect; this result may relate to reflexive knowledge depending closely on children's ability to recognize reflexive lexical items.

Questions Answered:

- (Q2) To what can the vocabulary effect observed by Lukyanenko, Conroy & Lidz (2014) be attributed: grammatical knowledge, processing speed, or experimental design?
- (A2) The effect in the NAME condition can be attributed to a task effect. The effect in the REFLEXIVE condition remains to be interpreted.

Lexical Access Speed Task

Figure 11 shows children's overall performance on the Lexical Access Speed task. Children were overwhelmingly successful in attending to the target object upon hearing the target word, as indicated by the sharp increase in the proportion of looking to

the target after 2300 ms, when the onset of the target word occurred (marked by the leftmost edge of the window). This looking behavior is significantly different from chance ($t(63)=17.58$, $p<0.001$). By 30 months, this type of task is exceedingly easy for children in general, but differences in the speed with which children orient to the target image are indicative of individual variation in speed of lexical processing.

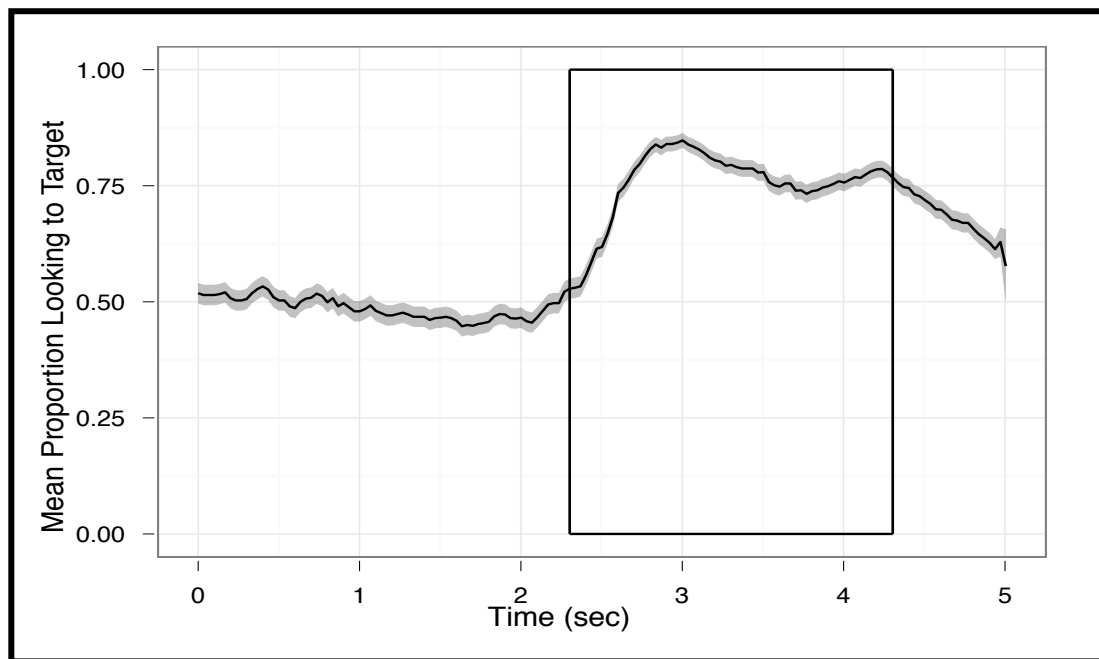


Figure 11 Looking behavior in Experiment 4 Lexical Access Speed Task. Outlined box represents critical window for analysis (2000ms following the onset of the disambiguating NP).

As noted above, the measure gathered from the LAS task was each child's mean latency to shift to the target image on distractor-initial trials¹⁷. LAS values ranged from

¹⁷ All response latency measures are taken relative to 300 ms following the raw disambiguation point, to account for the time required for saccade programming.

100.1 to 989.99 ms, with a median LAS of 329.496 ms (mean = 385.29)¹⁸; this suggests that at 30 months, children's speed of processing even at the lexical level is still quite varied.

Phrase Structure Integration Speed Task

Questions Addressed

(Q3) Do children successfully interpret hierarchical structure in the PSIS task?

Figures 12-13 present the results for the Phrase Structure Integration task by condition. Because the visual array in this case consists of a target item and two distractors, the graph in these cases shows the proportion looking to each item (rather than the binary distinction of looking to target vs. away from target to the distractor presented in previous figures). Because there are three possible items to attend to, chance is estimated at 33.3%, rather than 50%. Recall that in the SUPERLATIVE condition, where children hear sentences like *where's the biggest train*, the target item will be the large item, while in the SUPERLATIVE + ADJECTIVE condition, where children hear sentences like *where's the biggest red train*, the target will be the medium item (corresponding to the larger of the two similarly colored items). One obvious effect in both conditions is an overall bias to look at the largest item; looking to the largest item is significantly higher than chance in both conditions (SUPERLATIVE condition: $t(63)=16.1$, $p<0.001$;

¹⁸ Because the Principle C task is a between-subjects design, we computed separate median values for children who were run on each condition of the Principle C task. LAS values for children in the REFLEXIVE condition had a median of 333.67 ms (range = 100.1-989.99 ms, mean = 417.75 ms); LAS values for children in the name condition had a median of 300.3 ms (range = 122.36-934.27 ms, mean = 352.99 ms). LAS values did not differ significantly across conditions ($t(63)=-1.176$, $p>0.5$), so in the following analyses of the task, median split groups correspond to two separate median splits, one for each condition of the Principle C task.

SUPERLATIVE + ADJECTIVE condition: $t(63)=13.02$, $p<0.001$). For this reason, we compare performance in the window of analysis to pre-disambiguation looking behavior as a baseline. Behavior following the disambiguation point differs by condition. A 2 (condition) x 2 (probe- pre-disambiguation vs. critical window) ANOVA reveals a main effect of condition ($F(1,63)=111.3$, $p<0.001$), no main effect of probe ($F(1,63)=0.5487$, $p>0.1$), and a highly significant interaction between condition and probe ($F(1,63)=66.2$, $p<0.001$). Comparisons to investigate the interaction term show that looking to the largest item is significantly higher than pre-disambiguation performance in the SUPERLATIVE condition ($t(63)=-4.864$, $p<0.001$), but significantly lower in the SUPERLATIVE + ADJECTIVE condition ($t(63)=6.832$, $p<0.001$).

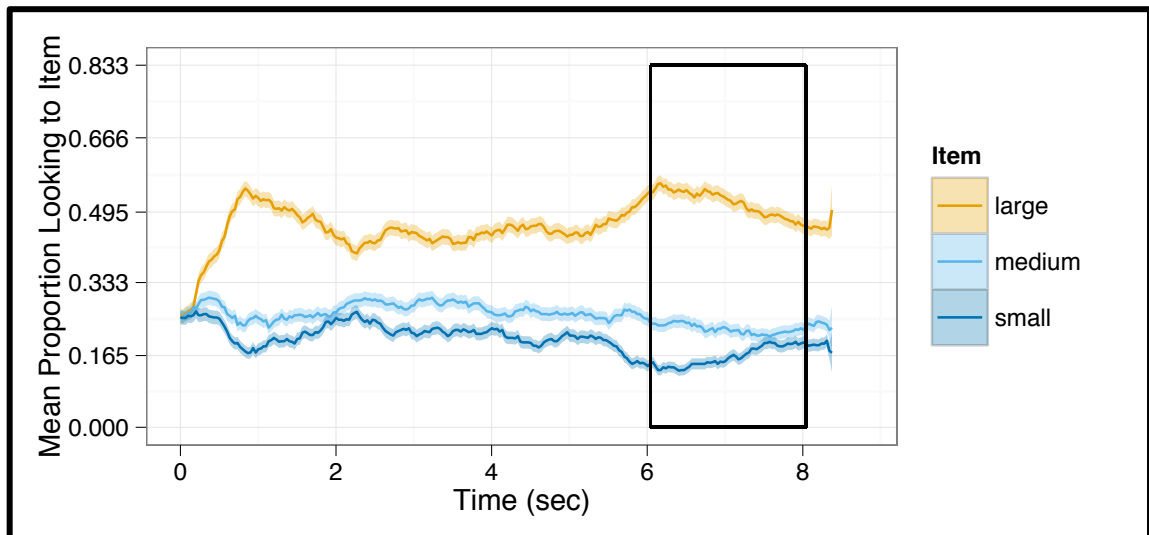


Figure 12 Looking behavior in Experiment 4 Phrase Structure Integration Speed Task SUPERLATIVE condition. Outlined box represents critical window for analysis (2000ms following disambiguation).

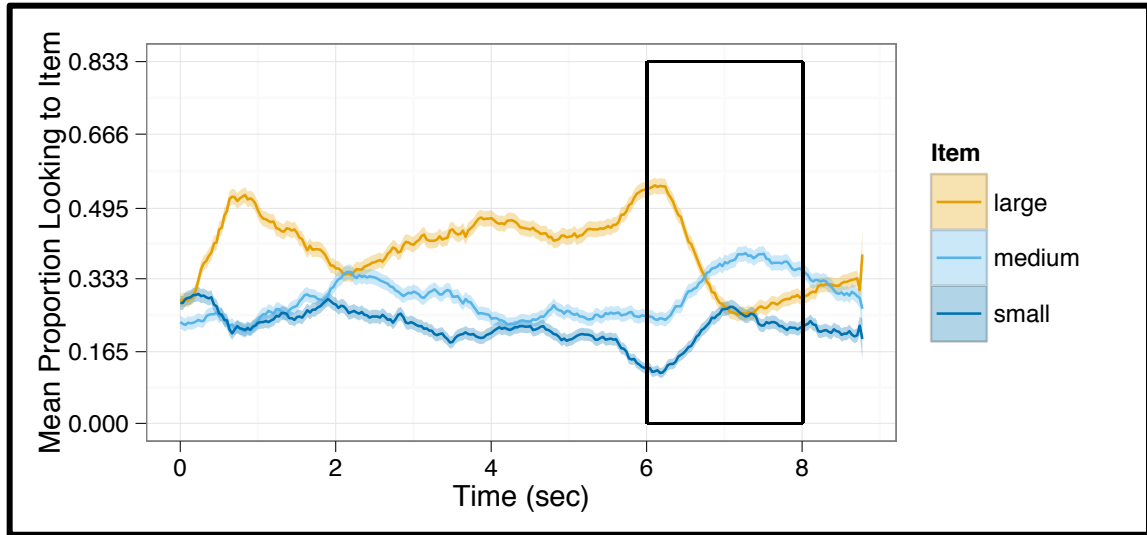


Figure 13 Looking behavior in Experiment 4 Phrase Structure Integration Speed Task SUPERLATIVE + ADJECTIVE condition. Outlined box represents critical window for analysis (2000ms following disambiguation).

As with the LAS task, the measures gathered from the PSIS task was each child's mean latency to shift to the target image on distractor-initial trials; a separate speed measure was gathered for each condition. PSIS SUPERLATIVE values ranged from 133.47 to 1868.53 ms, with a median PSIS (SUPERLATIVE) of 764.76 ms (mean = 791.12 ms).¹⁹ PSIS SUPERLATIVE + ADJECTIVE values ranged from 340.34 to 1985.32 ms, with a median PSIS (SUPERLATIVE + ADJECTIVE) of 1063.73 ms (mean = 1052.72 ms).²⁰

¹⁹ As with the LAS task, we calculated separate median values for each condition of the Principle C task (see footnote 16). PSIS (SUPERLATIVE) values showed a median of 798.13 ms in the REFLEXIVE condition (range = 133.47-1581.58 ms; mean = 819.15 ms) and 759.76 ms in the NAME condition (range = 225.23-1868.53 ms; mean = 763.43 ms). PSIS values did not differ significantly across conditions ($t(63)=-1.176$, $p>0.5$), so in the following analyses of the task, so we collapse across these multiple median split groups.

²⁰ PSIS SUPERLATIVE + ADJECTIVE values showed a median of 1061.06 ms in the REFLEXIVE condition (range = 500.5-1584.92 ms; mean = 1039.71 ms) and 1063.73 ms in the NAME condition (range = 340.34-1985.32 ms; mean = 1065.73 ms). PSIS values did not differ significantly across conditions ($t(63)=-1.176$, $p>0.5$), so in the following

In both conditions, we see increases in looking immediately preceding the disambiguation point (the target noun in the SUPERLATIVE condition and the target adjective in the SUPERLATIVE + ADJECTIVE condition, marked in these graphs by the left edge of the window), consistent with children's incremental interpretation of the adjective *biggest*. In the SUPERLATIVE condition, children primarily maintain looking to the largest item; this performance is consistent with accurate integration of *train* and successful interpretation of the target NP. In the SUPERLATIVE + ADJECTIVE condition, children show a sharp decrease in looking to the largest item and corresponding increased looking to the medium item; this performance is consistent with accurate integration of *red train* and successful interpretation (or revision of interpretation) of the target NP. In this task, accurate interpretation and selection of the target object requires children to treat the test sentences as hierarchically structured phrases. We see that children are able to select the target item in both conditions, suggesting that at 30 months, children's interpretations are hierarchically structured.

Questions Answered

(Q3) Do children successfully interpret hierarchical structure in the PSIS task?

(A3) **Yes**

Processing Speed Analyses: Comparison of Covariate Measures

Questions Addressed

(Q2) To what can the vocabulary effect observed by LCL be attributed: grammatical knowledge, processing speed, or experimental design?

analyses of the task, median split groups correspond to two separate median splits, one for each condition of the Principle C task

(Q4) Does our new measure of syntactic processing speed capture something distinct from the (lexical) processing speed measure formalized by Swingley et al (1999)?

From the LAS and PSIS tasks, we collected three measures of children's processing speed, at the lexical and syntactic levels. As noted above, MCDI information was also collected for each child. Figure 14 shows comparisons between each of these measures; Table 2 presents the correlation coefficients and significance terms. In sum, none of these four measures are significantly correlated with each other (all $p > 0.1$).

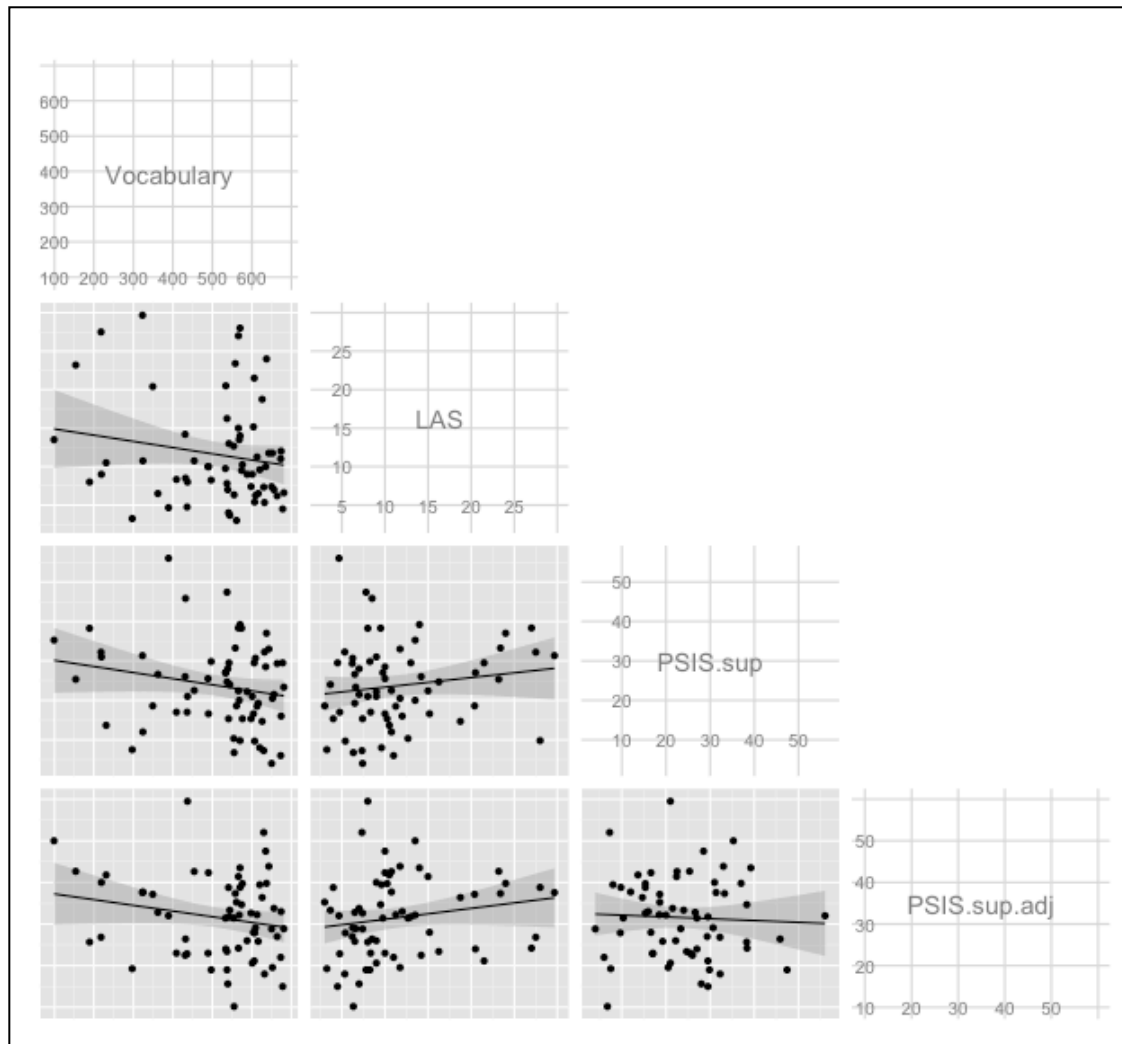


Figure 14 Comparisons of covariate measures in Experiment 4.

	LAS		PSIS (superlative)		PSIS (superlative + adj.)	
Vocabulary	-0.17	p=0.169	-0.2	p=0.104	-0.2	p=0.113
LAS			0.15	p=0.24	0.18	p=0.161
PSIS (superlative)					-0.05	p=0.709

Table 2 Correlations between covariate measures in Experiment 4

Previous research has also analyzed results by comparing performance of children in two groups, defined by falling above or below the median value (such as the high vocabulary vs. low vocabulary analysis presented by Lukyanenko, Conroy & Lidz). However, this analysis also fails to show any correlation between children who fall into the above-median or below-median groups. Table 3 presents the number of children whose values fall into each category across all measures; Table 4 presents the corresponding chi-squared tests. If any measure was correlated with another, we would expect children to fall predominantly into the boxes representing above-median on both measures or below-median on both measures (i.e. the shaded boxes in Table 3); however, the pattern we see is instead a roughly even distribution across all 4 possible boxes. Chi-squared tests show none of the median split groups for any measure to be related to the median split groups for any other measure (all $p > 0.1$), further confirming the results of the continuous measure analysis above.

	Fast LAS	Slow LAS	Fast PSIS (superlative)	Slow PSIS (superlative)	Fast PSIS (superlative + adjective)	Slow PSIS (superlative + adjective)
High Vocabulary	14	17	19	12	15	16
Low Vocabulary	15	18	13	20	17	16
Fast LAS			14	15	18	11
Slow LAS			18	17	14	21
Fast PSIS (superlative)					13	19
Slow PSIS (superlative)					19	13

Table 3 Comparison of covariate measures by median split groups

	LAS	PSIS (superlative)	PSIS (superlative + adjective)
Vocabulary	$\chi^2 = 0.0518$ p = 0.8199	$\chi^2 = 2.2522$ p = 0.1334	$\chi^2 = 0$ p = 1
LAS		$\chi^2 = 0$ p = 1	$\chi^2 = 2.27$ p = 0.1319
PSIS (superlative)			$\chi^2 = 1.5625$ p = 0.2113

Table 4 Chi-Squared distributions of median split groups. All $\chi^2(1, n=64)$

The lack of correspondence between any of these measures allows us to see more clearly two points about the nature of these processing speed measures, which have been suggested above. First, the lack of correlation between LAS and either measure of PSIS provides that the measures we have created for processing at the syntactic level are independent of processing at the lexical level. That is, interpretation in the PSIS task involves more than simply compounded lexical access, suggesting that children are in fact computing the structure of the sentence, or at least composing the meaning, in order to arrive at an interpretation.

Second, the lack of a relation between the two measures of PSIS themselves suggests that the processing required in each condition varies; we take this difference to be indicative of the potential additional requirement in the superlative + adjective condition for children incrementally parsing the sentence to revise their initial interpretation.

Further, two points of evidence here allow us to draw some interim conclusions related to our question of how vocabulary plays a role in children's interpretations in Principle C contexts. First, the lack of correlation between MCDI vocabulary and LAS suggests that vocabulary may not be related to processing speed at the lexical level at 30 months of age²¹. Second, the lack of correlation between MCDI vocabulary and either PSIS measure suggests that vocabulary may also be unrelated to processing speed at the syntactic level. Together, these points significantly weaken the argument that the vocabulary effect on children's Principle C knowledge is a reflex of variation in processing speed, as vocabulary seems to be independent of lexical or syntactic processing at this age.

Due to the lack of a relation shown between any of these measures, in the following analyses we treat each as an independent covariate measure.

²¹ Of course, it is possible that 30 months is simply beyond the age range when vocabulary can be reliably linked to lexical processing speed. This could be due to either ceiling effects on the vocabulary measure, or to the fact that there is less variability in lexical processing than there is at younger ages. Additionally, one could raise the methodological issue that tasks measuring processing speed in the literature generally use more trials; while our task included 8 trials, most tasks in previous research have used upwards of 20 trials. Ongoing research aims to further explore the connection between vocabulary and various measures of processing speed with increased number of trials to facilitate having enough distractor-initial trials to form reliable measures.

Questions Answered

- (Q2) To what can the vocabulary effect observed by Lukyanenko, Conroy & Lidz (2014) be attributed: grammatical knowledge, processing speed, or experimental design?
- (A2) The effect in the NAME condition can be attributed to a task effect. The effect in the REFLEXIVE condition remains to be interpreted; it cannot be attributed to variation in processing speed in any clear way.
- (Q4) Does our new measure of syntactic processing speed capture something distinct from the (lexical) processing speed measure formalized by Swingley et al (1999)?
- (Q4) **Yes.**

Growth Curve Analyses: Individual Differences in Speed of Interpretation

Questions Addressed

- (Q2) To what can the vocabulary effect observed by LCL be attributed: grammatical knowledge, processing speed, or experimental design?
- (Q5) Which of these factors, if any, predict performance in Principle C contexts? What inferences can we therefore make about the mechanism driving Principle C effects?

So far we have demonstrated that all children exhibit behavior consistent with knowledge of Principle C at 30 months. Our analysis of individual differences in processing seeks to determine what type of processing is implicated in this response, as a means to determine the underlying knowledge behind children's performance pattern. If speed of interpretation is predicted by processing at the syntactic level, then we can infer

that the underlying knowledge driving behavior is structural in nature; this dependency is predicted if children are using accurate knowledge of Principle C. Alternatively, if speed of interpretation is predicted by processing at the lexical level, then we can infer that the underlying knowledge driving behavior is non-structural in nature.

To examine speed of interpretation, we analyzed distractor-initial trials, as we did for the LAS and PSIS tasks in our generation of the processing speed measures. To ensure that response latency on distractor-initial trials is actually indicative of speed of interpretation of the test sentence, we compared performance to that of target-initial trials. If children rapidly shifted fixation equally often in both distractor-initial and target-initial trials, it would suggest that their fixation behavior is independent of interpretive processes and would not reflect a response to the target sentence, but rather some sort of behavior pattern technique that results from the nonlinguistic demands of the task and not to any property of the linguistic stimulus. Figures 15-16 present performance in each condition from the disambiguation point onwards comparing distractor-initial and target-initial trials. A two-way ANOVA over the NAME condition shows a main effect of onset (distractor- vs. target-initial) ($F(1,63)=19.39$, $p<0.001$), while the corresponding ANOVA over the REFLEXIVE condition shows no such effect ($F(1,63)=2.414$, $p>0.1$).

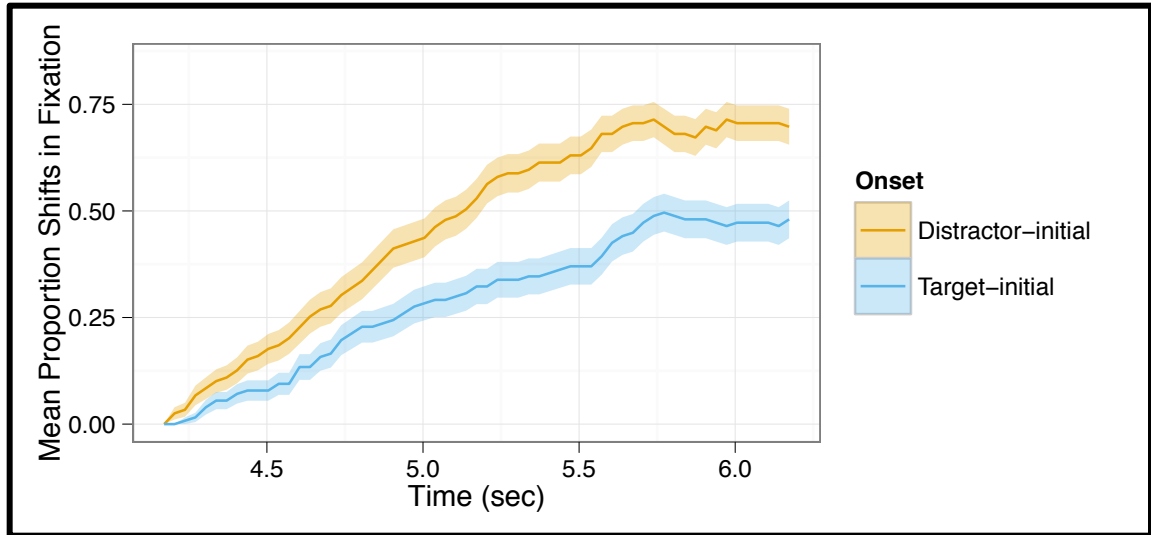


Figure 15 Onset-contingent looking behavior in Experiment 4 Principle C Task NAME condition. Performance measured from the shifted disambiguation point.

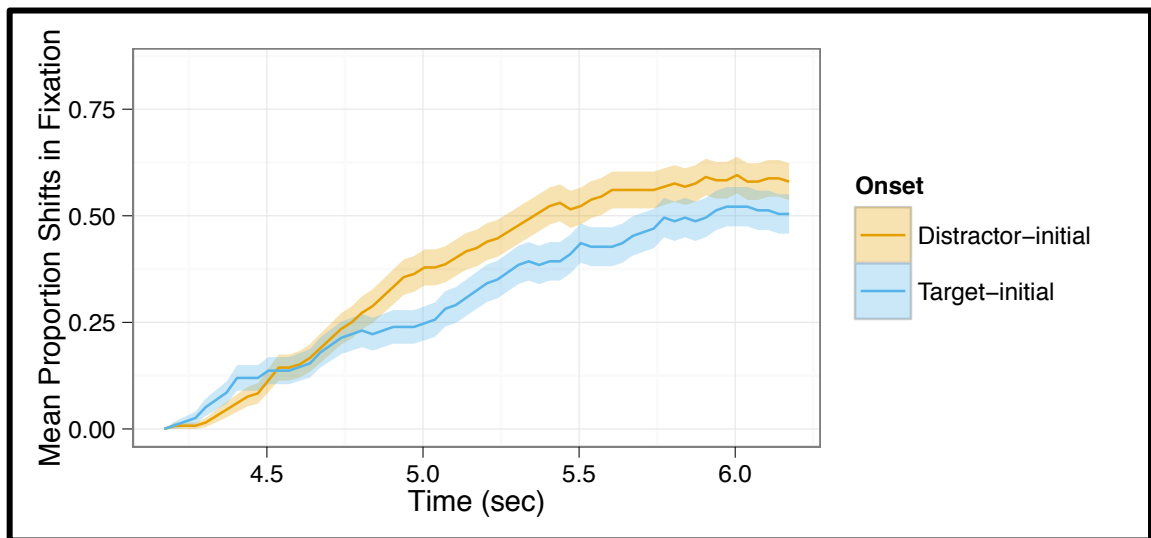


Figure 16 Onset-contingent looking behavior in Experiment 4 Principle C Task REFLEXIVE condition. Performance measured from the shifted disambiguation point.

This result suggests that the response latency value is not an effective measure of speed of interpretation in the REFLEXIVE condition, where the response is less closely time-locked to the linguistic stimulus. We discuss a number of possibilities to account for this difference in Section 4.4. Because the REFLEXIVE condition does not present a

situation in which response latency is an effective measure of speed of interpretation, we therefore restrict our analysis to performance in the NAME condition.

In order to most closely capture children's response latency in the NAME condition trials, corresponding to their first shift in fixation from the distractor to the target, we restricted our analysis to the first 1500 ms post-disambiguation, as this is the window when the majority (over 85%) of first fixation shifts occur. We used growth curve analysis over the NAME condition distractor-initial trials in order to compare performance of individuals based on their values on the covariate measures (MCDI vocabulary, LAS, PSIS (SUPERLATIVE), and PSIS (SUPERLATIVE + ADJECTIVE)). Looking behavior was modeled with fourth-order (quartic) orthogonal polynomials, fixed effects for each covariate measure (treated as continuous measures, each in a separate model), and random effects of participant on all time terms. Table 5 presents an analysis of model fit for each set of models. The deviance statistic -2LL provides a measure of model fit (with smaller values representing a better fit). Change in deviance is distributed as chi-square, with degrees of freedom equal to the number of parameters added to the model. The base model represents a model with no effect of covariate measures. When we observe an effect on one of the polynomial terms in any of the covariate measure models, this can be interpreted as significant improvement in model fit over the base model when that covariate measure is incorporated into the model. The polynomial term that bears the effect is indicative of the component of the looking behavior that is captured by integrating the covariate measure into the model. (I will address each effect we observe and how the effect might be interpreted in further detail below.) There was a significant effect of PSIS (SUPERLATIVE + ADJECTIVE) on the intercept term (Estimate = -0.002023,

SE = 0.0008, $p < 0.05$), indicative of overall greater proportion of looking to the non-reflexive image in this time window by children with faster PSIS (SUPERLATIVE + ADJECTIVE). Additionally, there was a marginally significant effect of PSIS (SUPERLATIVE) on the linear term (Estimate = -0.01973, SE = 0.01, $p = 0.0556$), and a marginally significant effect of vocabulary on the quadratic term (Estimate = -0.00118, SE = 0.00058, $p = 0.0506$); effects on these terms are indicative of differences in speed to shift attention to the non-reflexive item. All other effects were non-significant. Model fits for each covariate measure model are shown in Figure 17.

Model	-2LL	Δ LL	p
Base	-2648	---	---
Vocabulary			
intercept	-2650	2	n.s.
linear	-2652	4	n.s.
quadratic	-2655	7	0.0506
cubic	-2656	8	n.s.
quartic	-2656	8	n.s.
Lexical Access Speed			
intercept	-2649	1	n.s.
linear	-2649	1	n.s.
quadratic	-2650	2	n.s.
cubic	-2650	2	n.s.
quartic	-2651	3	n.s.
PSIS (superlative)			
intercept	-2649	1	n.s.
linear	-2652	4	0.0556
quadratic	-2653	5	n.s.
cubic	-2653	5	n.s.
quartic	-2653	5	n.s.
PSIS (superlative + adj.)			
intercept	-2654	6	0.0172
linear	-2654	6	n.s.
quadratic	-2655	7	n.s.
cubic	-2655	7	n.s.
quartic	-2655	7	n.s.

Table 5 Model fit results of Growth Curve Analysis by covariate measure

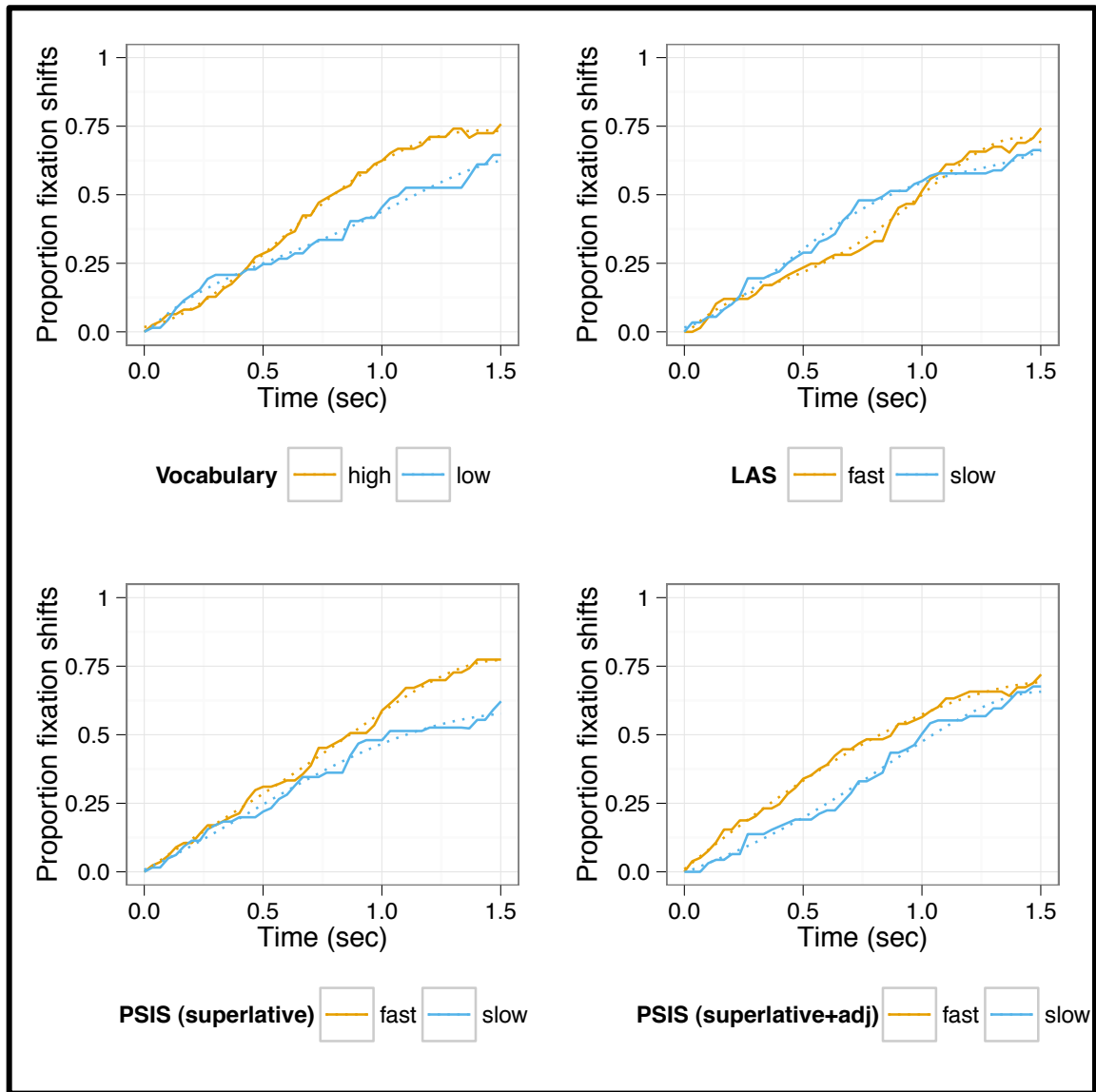


Figure 17 Model fits to observed data for Experiment 4 Principle C Task. Observed data are represented by solid lines. Model fits are represented by dotted lines.

These results show that by 30 months, speed of interpretation in Principle C contexts does not seem to be directly related to lexical processing speed, but is related to

at least one measure of syntactic processing speed²². Differences in the linear component of performance, as indicated by an effect on the linear term of the model based in PSIS (SUPERLATIVE) differences, are a clear indication of differences in speed of interpretation. This finding is consistent with the prediction that if reaching the correct interpretation in the Principle C condition requires structure building, then performance in that condition would be linked to children's speed of processing syntactic information. This finding thus suggests that by 30 months, children's interpretation in Principle C contexts is driven by syntactic knowledge.

An effect on the intercept term, as seen with the model including the effect of PSIS (SUPERLATIVE + ADJECTIVE), is not directly indicative of a difference in speed (generally indicated by the rate of change, or slope). Rather, it indicates a difference in proportion of fixation shifts to the target which is constant across the entire time window; this is effectively the proportional measure derived from standard analyses of timecourse data which collapses across windows of time. That is, an effect on the intercept term is generally interpreted as corresponding to a difference in 'accuracy' (defined as fixation to the target image) rather than a difference in response speed. Because our analysis here encompasses only the first 1500 ms post-disambiguation, this is not an ideal measure to accurately estimate overall accuracy in this task. One might, however, consider one way in which even the intercept term, which does not take into account changes across the time window of analysis, might still be reflective of a difference in speed. Recall that the

²² Recall that we predict that the PSIS measure gained from SUPERLATIVE + ADJECTIVE trials may not be an effective measure of syntactic processing speed, as it potentially confounds two processes required for interpretation- integration of information into the structure, and revision of initial incorrect parses. Thus the lack of effect on a model term indicative of slope differences is in this case not surprising.

dependent measure for these analyses is the proportion of *shifts* in fixation (from the distractor to the target image). Additionally, this analysis has been measured exclusively over distractor-initial trials, in order to accurately assess shifts in fixation to the target image. In this subset of the data, behavior at time point 0 is consistent across the dataset—no shift in fixation has yet occurred on any trial. Given that all behavior is identical at the beginning of the window of analysis, a difference which is stable across the majority of the time window, being realized as an effect on the intercept term, could arise from a slight difference in slope at the beginning of the window that then levels off. In this way, an effect on the intercept term could even be interpreted as resulting from a difference in speed of response. Thus one might even interpret the effect seen in the PSIS (SUPERLATIVE + ADJECTIVE) analysis to reflect differences in speed of interpretation in Principle C contexts.

We additionally found an effect on the quadratic term in the model examining the effect of vocabulary. An effect on this term also corresponds to differences in slope, but additionally captures variation in slope across the time window. While we did not predict vocabulary to affect speed of interpretation, it remains an open question what vocabulary may be indexing. It is possible that this effect arises because vocabulary is in part a consequence of syntactic knowledge, but in some way that is not captured by syntactic processing. It could be that children with better syntactic abilities in general have larger vocabularies, which in turn could yield an indirect effect on interpretation here. We discuss further possibilities with respect to the role of vocabulary size in Section 4.4.

Questions Answered

- (Q2) To what can the vocabulary effect observed by LCL be attributed: grammatical knowledge, processing speed, or experimental design?
- (A2) The effect in the NAME condition can be attributed to a task effect. The effect in the REFLEXIVE condition remains to be interpreted; it cannot be attributed to variation in processing speed in any clear way. However, given that vocabulary predicts speed of interpretation in Principle C contexts, it may be that vocabulary is related to a more global (i.e. combined) measure of syntactic competence or processing.
- (Q5) Which of these factors, if any, predict performance in Principle C contexts? What inferences can we therefore make about the mechanism driving Principle C effects?
- (A5) Syntactic processing predicts speed of interpretation in Principle C contexts. This effect is expected only if interpretation in Principle C contexts is dependent on structure building. MCDI Vocabulary size also affects speed of interpretation in Principle C contexts; the implications of this finding remain undetermined.

4.4 General Discussion

Summary of findings

Using a 2000ms window of analysis, we find:

- (A) that children at 30 months show success in interpreting both Principle C and reflexive sentences;

(B) no effect of vocabulary on overall Principle C performance, unlike Lukyanenko, Conroy & Lidz (2014);

(C) an effect of vocabulary on reflexive sentences, replicating the finding of LCL.

In the analysis of covariate measures, we find:

(A) no correlations between our measures of vocabulary, lexical processing and syntactic processing;

(B) that response latency is not a useful measure for assessing interpretation in reflexive contexts;

(C) that syntactic processing but not lexical processing predicts speed of interpretation in Principle C contexts.

The research presented here shows several key findings. We replicate results of LCL, showing that by 30 months, children exhibit behavior consistent with adult-like knowledge of Principle C. We find a vocabulary effect, as Lukyanenko et al. did, however while vocabulary predicted performance in both conditions in their experiment, vocabulary only affected overall performance in the REFLEXIVE condition in our study. This difference is most likely related to differences in task design, in utilizing a between-subjects design rather than within-subjects. Limiting comprehension to one sentence type may have made it easier for low vocabulary children in our study to interpret the structures of the test sentences. This suggests that at least in Principle C contexts, to vocabulary effect observed by LCL was likely due to extralinguistic factors; it may be these extralinguistic factors that contribute to word learning, which manifests as differences in vocabulary.

Components of our analysis are novel in two ways. First, we have developed a method for measuring processing speed at the syntactic level. Previous methods for assessing processing speed have targeted lexical information; our new syntactic processing speed task requires children to interpret a sentence compositionally, reflecting hierarchical syntactic structure. We demonstrate that the measures gained from this task are independent of lexical processing, suggesting that they successfully index speed of processing syntactic information.

Most importantly, our analysis is novel in using individual differences at the level of deployment to suggest underlying similarities at the level of knowledge. Because we see a clear contribution of structure-building processes in predicting the time-course of success in the Principle C task, we can conclude that this success is driven by structural knowledge. Moreover, because speed of response in the Principle C task is not predicted by the speed of non-structural processes like lexical access, we gain further support for the role of structure in explaining children's behavior. Finally, if children's success with Principle C sentences were driven by non-structural heuristics for reaching adult-like interpretations, we would not expect differences in the speed of interpreting Principle C sentences to be correlated with the speed of building structure. Taken together, then, these findings suggest that children's interpretations in these experiments are driven by Principle C, suggesting that this knowledge is in place at least by 30 months of age.

Open Questions

One question that remains to be answered is what the role of vocabulary in demonstration of Principle C knowledge may be. Because our between subjects-design

failed to reveal an effect of vocabulary on performance, but the within-subjects design of LCL did show such an effect, it seems likely that the earlier vocabulary effect is likely due to nonlinguistic contributors to performance, rather than relating to grammatical development. Additionally, however, we see that vocabulary is predictive of speed of interpretation in Principle C contexts, but it does not correlate with any of our measures of processing speed. Thus exactly what contribution vocabulary makes to speed of interpretation remains unclear. It seems that MCDI vocabulary may index a simpler factor in children's word-learning capabilities. We consider two possibilities: first, the vocabulary effect could be an effect not of children's full vocabulary, but of their knowledge of the particular verbs used in these test sentences. However, we would expect this to affect performance in both conditions equally; instead, we see a vocabulary effect only in the REFLEXIVE condition. Alternatively, the effect could be evidence of children with smaller vocabularies being less adept at learning new words, making it more difficult for these children to process the names and relevant reference relations in these sentences. Again, however, it is unclear how this would manifest in one condition but not the other. More generally, vocabulary could be a type of composite measure which is indirectly tied to multiple aspects of grammatical and cognitive development. Essentially, this possibility is consistent with vocabulary being related to grammatical development without directly implicating any one specific factor of grammatical knowledge or grammatical processes. This possibility is appealing in light of previous research showing vocabulary's relation to various measures of grammatical development, and given that our growth curve analysis found vocabulary to be predictive of speed of response in Principle C contexts. However, it is also somewhat loosely defined; closer

investigation is necessary to determine which aspects of grammatical development vocabulary is most likely to affect, as well as how alternative measures of grammatical development might be generated, in order to compare with vocabulary.

An additional question relates to the differing timecourse of response in the REFLEXIVE condition as compared to the NAME condition. Behavior is closely time-locked to the linguistic stimulus in the NAME condition, such that children shift their attention to the target item within the first 2000 ms after hearing disambiguating information in the test sentence (in this case, the target direct object NP). Comparatively, behavior in response to linguistic stimulus is delayed in the REFLEXIVE condition, only appearing about 4000 ms after disambiguation. This pattern may relate to the nature of the constraints Principle C and Principle A, and differences in how reference is resolved in these differing contexts. The resolution of a pronoun's referent in a Principle C context is a forward-looking process; in other words, a listener can predict that NPs in the c-command domain of a pronoun will be disjoint in reference (Kazanina et al., 2006; c.f. discussion in Conroy et al., 2009). Contrastively, reference resolution for a reflexive requires accessing the antecedent from memory, in a backward-looking search (Sturt 1999; Dillon 2010). Thus failure to predict performance in reflexive contexts may be a result of the antecedent retrieval process not being adequately indexed by eye movements in this task.

4.5 Conclusion

The research presented in this chapter sought to identify the specific character of knowledge which drives performance by examining the processes required to deploy this

knowledge in real time. I identified three possible factors which could contribute to interpretation in Principle C contexts: vocabulary, processing of lexical information, and processing of syntactic information. I show that speed of interpretation in Principle C contexts (performance) is predicted by processing speed at the syntactic level but not at the lexical level (deployment). This dependency between syntactic processing and performance suggests underlying knowledge which is based in phrase structure and the ability to perform computations over such structure (knowledge). The success of this novel analytic approach opens the door to a new way of assessing many areas of syntactic knowledge.

5 Conclusion

A primary challenge in acquisition research is that while we aim to diagnose children's underlying competence, we only have access to their performance. Beyond diagnosing simply whether competence is adult-like or not, it is also important to identify how this competence comes into place. Can we identify a pattern of children successively upgrading interpretive mechanisms that approximate the adult grammar before adult-like grammatical knowledge is available? Or is the pattern instead all-or-nothing, when children's earliest performance consistent with adult grammatical knowledge seems to also be driven by adult-like competence? Understanding the developmental trajectory for a given linguistic phenomena is important to our understanding of how such a phenomena arises in the grammar. In this dissertation, I explored a case study of children's early syntactic knowledge. My in-depth analysis of Principle C at 30 months has provided novel insights into diagnostics for underlying competence by utilizing two distinct methods of analysis which probe both children's performance as well as the deployment processes required to implement underlying knowledge in performance. Additionally, this work has served to expand our understanding of the developmental trajectory of Principle C, which can contribute to debates about the origin of this constraint as part of the grammar.

This dissertation has explored the developmental trajectory of Principle C through two primary analytic approaches. First, I explored children's behavioral response across multiple linguistic contexts in which Principle C applies. This research demonstrated that

across contexts, other proposed interpretive mechanisms based in non-structural information cannot account for all of children's behavior. Second, I used individual measures of processing as a predictor of performance. Because different interpretive strategies will require different processes to be deployed, I was able to use measures of these deployment processes as a diagnostic rather than a flaw in children's performance to be controlled for. This research showed that syntactic processing predicts performance, implicating structural composition as a mechanism used in children's early interpretations in Principle C contexts. Together, these findings have not only demonstrated new methods for interpreting children's performance, but have also extended our knowledge of the developmental pattern that characterizes Principle C.

In this conclusion, I will summarize the key findings presented in this dissertation, as well as their broader implications, by addressing and answering the questions originally posed in the Introduction.

5.1 Key Findings

Question 1: At what point in development are Principle C effects observable in children's behavior?

Chapter 2 provided insight into the age at which Principle C effects emerge. Identifying the onset of behavior consistent with adults is critical to expanding our understanding of the developmental pattern of Principle C. This is the first step towards identifying the point in development at which we believe Principle C knowledge to be fully acquired. It is this point that can be taken as the end-point for whatever development occurs. It remains to be determined whether this development is characterized by growth

of knowledge through representational change or by recognition of knowledge through identifying some mapping between strings in the language and available representations. In order to answer this more complex debate, it will first be necessary to identify the period during which this development occurs, starting with the point at which it is completed. Previous research showed Principle C effects in children as young as 30 months. In Chapter 2, I explored performance in 24 and 30 month-olds (Experiments 1 and 2), showing that children exhibit a preference for a disjoint interpretation at 30 months, but exhibit no such preference at 24 months. This finding suggests that the earliest age at which we can observe Principle C effects in children is somewhere between 25 and 30 months. It may be noted here that even this finding does not fully rule out Principle C knowledge at early ages. It is possible that 24 month-olds do have the underlying knowledge available but are not able to display that knowledge within the context of our task. In fact, children could potentially have adult-like knowledge, or even an alternative interpretive mechanism such as those discussed in Chapters 2 and 3; but if this is the case then at this point in development they are not able to exhibit that knowledge outwardly in a way that we can measure. The ability to deploy whatever knowledge drives early Principle C effects seems to arise sometime between 25-30 months. Now that we have an idea of when Principle C effects emerge in the developmental timeline, we can begin to explore whether these are actually caused by Principle C. That is, we will investigate whether children's early interpretations in Principle C contexts could be driven by some other non-structural interpretive strategies.

Question 2: Are early Principle C effects attributable to knowledge of Principle C, or does this early behavior arise from an alternative interpretive mechanism?

We have now established a potential “end point” of acquisition of Principle C as 30 months; this is the earliest age at which we are able to identify the preference for disjoint interpretations consistent with knowledge of Principle C. Now that we have identified the onset of adult-like behavior, we want to probe what type of knowledge drives this early behavior. Do children really know Principle C at 30 months? Or are they using some “cheating strategy”—an interpretive mechanism based in some non-structural knowledge. Chapters 2 and 3 identified two possible interpretive mechanisms which have been proposed as alternatives to Principle C in young children to account for early Principle C effects. Chapter 2 began with the possibility that children could use transitive syntax as a cue to a disjoint interpretation. Research from the literature on syntactic bootstrapping has robustly shown that children interpret sentences with transitive syntax as corresponding to events with two participants. Because initial studies used the basic transitive *X VERBs Y* structure, it is possible that children could employ this bias to interpret reference as an initial strategy rather than Principle C. To identify which mechanism is responsible for children’s behavior, I identified a linguistic context where these mechanisms predict differing performance. Experiments 1 and 2 investigated children’s interpretations of sentences such as (81) and (82).

(81) She’s patting Katie’s head.

(82) She’s patting her head.

Both of our specifications of a transitivity bias are formalized as relations between transitive syntactic frames and a particular interpretation. These transitivity biases only predict different interpretations for transitive and intransitive structures; they do not take into account the content of the NP arguments, only the number of arguments. By these biases sentences (81) and (82), both being transitive clauses, should be interpreted in the same manner and interpretations should not vary across sentence types. Unlike these transitivity biases, Principle C takes into account the type of nominal element in the argument NPs. Principle C predicts only one possible interpretation for (81), but two possible interpretations for (82), and thus expects a divergence in performance across sentence types. Experiment 1 showed that at 30 months, children's behavioral pattern differs by sentence type. This performance is consistent with knowledge of Principle C, and cannot be accounted for by a bias based in transitive syntax.

Chapter 3 explored the possibility that children could use linear precedence of a pronoun over an R-expression as a cue for a disjoint interpretation. Because we know that children use linear order as an initial interpretive strategy before they fully comprehend sentence structure, it is possible that are utilizing this strategy rather than syntactic knowledge in their early interpretation of reference relations. This study investigated children's interpretations of sentences such as (83) and (84).

(83) She's painting the house that's in Katie's lap.

(84) The house that she's painting is in Katie's lap.

A linear bias predicts that interpretations should not vary across sentence types, as the linear order of the pronoun preceding the R-expression is maintained across sentence types. Because c-command of the pronoun over the R-expression obtains in (83) but not (84) Principle C predicts only one possible interpretation for (83), but two possible interpretations for (84), and thus expects a divergence in performance across sentence types. Experiment 3 showed that at 30 months, children's response pattern is consistent with knowledge of Principle C, and cannot be explained by a bias based in linear order.

Together, these experiments show that across all tested contexts, Principle C is the only interpretive mechanism which is able to account for behavior in all cases. While these results do not conclusively identify Principle C as 30 month-olds' interpretive mechanism, they definitely serve to narrow the field of possible alternatives. These findings also serve to expand the linguistic contexts in which 30 month-olds exhibit behavior consistent with knowledge of Principle C.

Question 3: What are the processing mechanisms through which children's knowledge is deployed?

One monumental challenge to investigating children's competence is that it is mediated by performance. Further, performance is dependent not only on children's knowledge, but also on any processes required to implement that knowledge. Rather than attempting to filter out the effects of these processes, we have instead taken advantage of these processes as a means of diagnosing the underlying knowledge requiring these processes. In Chapter 4 I compared the predictiveness of different measures of processing speed. Because we know that interpretation via Principle C will require syntactic

composition, we expect that speed of interpretation in a Principle C task should be dependent on a child's speed of processing syntactic information. Experiment 4 showed that individual variance in syntactic but not lexical processing speed predicts individual variance in the speed of interpretation in Principle C contexts. This dependency suggests that whatever strategy is responsible for interpretation in Principle C contexts must incorporate syntactic structure building. This result serves to once again narrow the realm of possible alternative interpretive strategies to account for early Principle C effects; because performance is predicted by syntactic processing speed, the interpretive strategy must be structural in nature.

Question 4: How did children arrive at the knowledge state that drives their early Principle C effects? Can we contribute to the innate vs. learned debate?

Constraints on interpretation pose an interesting learning problem, as there is inherently no information in the child language learner's input which gives explicit positive evidence for such a property. Computational models can be implemented to show that learning from indirect negative evidence is possible (e.g. Regier & Gahl, 2004; Pearl & Lidz, 2009; which provide accounts for the findings of Lidz, Waxman & Freedman, 2003). The first task, however, is to define the space in which learning might take place, and identify what evidence children might use. While I do not attempt to provide an ultimate answer to the innate vs. learned question for Principle C, the research presented here does take several steps toward identifying a solution.

First, as discussed above, this dissertation has narrowed in on defining the end point of learning. These results suggest the earliest age at which Principle C effects are

present in children's performance. Further, the elimination of the most likely alternative interpretive mechanisms show that we do not find any point at which we can confirm that children use a non-adult interpretive strategy in Principle C contexts. We implicate syntactic processing in interpretation in Principle C contexts; this finding eliminates the possibility that interpretation is based in any non-structural mechanism. While these results still do not concretely identify Principle C as the driving force behind these earliest Principle C effects, they have significantly narrowed the field of possibility for viable alternatives.

Second, I offer a brief analysis of the input that children might use to recognize Principle C if a learning account were implemented. Consider the set of sentences in (85-89).

- (85) **a. He₁ saw [NP John₂]**
 b. John₁ saw [NP himself₁]
- (86) **a. He₁ saw [NP John's₂ mother]**
 b. John₁ saw [NP his_{1/2} mother]
- (87) a. [NP His₁ mother] saw John_{1/2}
 b. [NP John's₁ mother] saw him_{1/2}
- (88) **a. He₁ thought [CP that John₂ was being nice]**
 b. John₁ thought [CP that he_{1/2} was being nice]
- (89) a. [CP That he₁ was being nice] pleased John_{1/2}
 b. [CP That John₁ was being nice] pleased him_{1/2}

These sentences constitute a sampling of the types of sentences that children might encounter. While all of the sentences contain a pronoun and an R-expression, only three of the ten are such that the pronoun c-commands the R-expression (85a, 86a, and 88a). Thus it seems clear already that the subset of sentences over which Principle C will even apply is not necessarily a large portion of the potential input. Importantly, however, it is exactly this type of sentence which children would need to receive as input in order to learn Principle C. That is, children would need to recognize that when a pronoun c-commands an r-expression, the interpretation of such sentences is never such that the pronoun co-refers with the r-expression.

In order to examine the availability of input relevant to the acquisition of Principle C, I compared data from two corpora in CHILDES, a databank of transcribed corpora of child-directed speech databank (MacWhinney, 2000): the Home-School Study of Language and Literacy Development corpus, and the NewEngland corpus, both of which are composed of transcribed interactions between parents and children. Together, they comprise data from input to 116 children ages 2;3-5;0, with a total of 26,210 parent-to-child utterances.

In my search of these corpora, I focused on three types of sentences. First, as we have noted that even with innately specified knowledge of Principle C it will be necessary to determine which lexical items act as pronouns in a given language, I tallied the number of utterances that contained pronouns. As Principle C will most naturally apply over sentences with 3rd person singular pronouns, I restricted my search to this subset of pronouns²³. Next, from this set of sentences, I recorded the subset which also

²³ Due to the first-pass nature of this analysis, I also excluded *it* from these counts.

contained an R-expression. Finally, I tallied the subset within this set of sentences where the R-expression was c-commanded by the pronoun. Recall that it is this final sentence type which is critical for a learning theory of Principle C where the set of considered hypotheses is not innately restricted.

Table 6 shows the results of the corpora search, in terms of the number of occurrences of the relevant sentence types as well as the proportion of the corpora that they represent. Even the simplest sentence type, considering all sentences in which a 3rd person pronoun is used, comprises under 5% of the input. Sentences with both a 3rd person pronoun and an r-expression occurred in less than 1% of the input in both corpora, and of these, sentences where the pronoun c-commanded the r-expression made up less than one tenth of a percent of children's input.

Sentence Type	Input Available	
Pronoun (he, she, him, her)	983	3.75%
Pronoun + R-expression (any order)	28	0.11%
Pronoun + c-commanded R-expression	11	0.04%

Table 6 Frequency information from two corpora. Numbers reported: occurrences of parent-to-child utterances with the relevant element(s); proportion of corpus these occurrences represent.

Additionally, across both corpora, I found no sentences exhibiting backwards anaphora (cases where a pronoun linearly precedes but does not c-command an R-expression)²⁴. Thus even if children were able to take advantage of the rare cases of the critical sentence type, the sentence type which would disambiguate the correct Principle

²⁴ Thus sentences which contained both a pronoun and an r-expression where the pronoun did not c-command the R-expression were cases where the R-expression preceded the pronoun linearly.

C generalization from a linear constraint is not present in the input. This point is further evidence, then, that the available evidence from which children would be able to generalize Principle C is extremely rare.

One assumption we have made thus far is in thinking that 1% of the input available to the child is negligible enough that learning from it is not plausible. But is 1% really that small? First, consider the total amount of speech input that children might be exposed to. Cameron-Faulkner, Lieven & Tomasello (2003) analyzed interactions between 12 parent-child dyads from the Manchester corpus (Theakston et al., 2001) in the CHILDES database. They estimate that children are exposed to approximately 7,000 utterances in the course of a day. Note, however, that input estimates often vary widely. Hart & Risley (1995) showed that Socio-Economic Status (SES) has a serious impact on the quantity of language input that children are likely to receive. They estimate that children from high-SES families hear about 11,000 utterances per day, while children from low-SES families may only hear around 700 utterances per day.

Once we are able to estimate children's input, we need to estimate the duration of time across development when children's input would be relevant to learning Principle C. This will allow us to estimate the total amount of utterances that a child might hear which would provide relevant input to generalizing Principle C. As discussed above, the research presented in this dissertation suggests that 30 months is a likely end point to the acquisition process. All available evidence suggests that behavior at 30 months is consistent with knowledge of Principle C; further, no alternative interpretive mechanism yet studied has been able to account for children's performance at this age. We can therefore limit ourselves to the input that children receive by 30 months old.

We have paid comparatively little attention to when the starting point of acquisition might be. We certainly don't want to start the clock at birth; children must minimally have the building blocks required to encode the speech input in a meaningful way for solving this problem. That is, children are not likely to be gathering data about syntactic phenomena before they are able to accurately segment words, or in the case of Binding Theory, before they have an accurate grasp of which words of their language correspond to pronouns, anaphors, and R-expressions. So when do children show early comprehension of pronouns? Children begin to produce pronouns, starting with first person, by 15 months (Macnamara, 1982). However, the time course for comprehension of third person pronouns, which are most relevant for the Principle C context sentences we are concerned with, remains unresolved. Several studies have shown that both comprehension and production of third person pronouns lag behind those of first and second person pronouns, showing limited comprehension even at 30 months (Charney, 1980; Strayer, 1977). These early results contrast with the research presented herein, which shows that children have enough of a grasp of pronouns to succeed in interpreting Principle C sentences by this age. The present scarcity of evidence to accurately pinpoint the onset of pronoun knowledge in children's grammar limits us from estimating when accurate encoding for Principle C might be in place.

Another component of knowledge children will need for speech input to be effectively encoded is syntactic comprehension in order to accurately compute the relevant c-command relations. Although we don't have a direct estimate of when children are able to compute c-command relations, we could take as a start the age when children begin to use syntactic structure to interpret sentences. Previous research suggests that

this time point may fall between 21 and 24 months. As discussed in Section 3.2, Gertner & Fisher (2012) showed that 21 month-olds rely on order of NPs to interpret sentences before they have access to full sentence structure. They found that 21 month-olds interpret conjoined subject intransitive sentences like (90) as having the same interpretation as a transitive structure like (91).

(90) The boy and the girl are gorging!

(91) The boy is gorging the girl!

By 24 months, children no longer use this word order strategy, but rather rely on syntactic cues to interpret similar sentences (Naigles, 1990). Based on these findings, it seems that the ability to compute syntactic relations to interpret sentences arises around 24 months of age. If we take this time point as the rough starting point when children have available the necessary component knowledge to reliably encode the input relevant to Principle C, it would mean that any learning that might take place occurs in roughly a six month period, from 24 to 30 months.

If we take these estimates of the number of sentences a child may hear in a day and the time period when these sentences could be encoded appropriately, we can now estimate the raw amount of relevant input children may receive from which to generalize Principle C. Six months of 7,000 utterances a day would stack up to 1,277,500 utterances. How many of these sentences would actually help a child generalize a constraint on interpretation like Principle C? In the corpus analysis presented above, sentences which contained a pronoun that c-commanded an R-expression made up only 11 of 26,210

utterances²⁵. With these two sets of numbers, we can now estimate that children might see approximately 536 data points relevant to Principle C. In real numbers, our original estimate of 1% of the input seems less negligible.

One way to determine whether this level of data available in the input is enough to learn from is by comparison to similarly frequent (or similarly rare) items. We can investigate the age of acquisition for words of similar frequency to determine whether comprehension at 30 months with this limited input is really plausible. The SUBTLEX_{US} corpus is comprised of 51.0 million words transcribed from American film subtitles (Brysbaert & New, 2009). A sample of words in this corpus that have similarly low frequencies comparable to our estimated Principle C-relevant input (420 per million, around 21,400 occurrences in this 51 million-word corpus) are presented in Table 7. To determine whether words of this frequency have been acquired by children by 30 months, I included the proportion of 30 month-olds that are reported to produce this word (as reported in the MCDI Lexical Norms Database).

²⁵ Recall, however, that while this sentence type potentially exhibits Principle C constraints, it is also consistent with alternative interpretive hypotheses, such as a Linear Bias. Backward anaphoric sentences where a pronoun precedes but does not c-command an R-expression were entirely absent from this dataset.

Word	Raw Frequency	Frequency Per Million	Reported Production at 30 Months
<i>mom</i>	21950	430.39	98.8%
<i>friend</i>	21384	419.29	81.3%
<i>move</i>	21325	418.14	n/a
<i>same</i>	21276	417.18	45.0%
<i>job</i>	21063	413.00	n/a
<i>tonight</i>	21047	412.69	52.5%
<i>went</i>	20987	411.51	n/a
<i>son</i>	20949	410.76	n/a

Table 7 Frequency information from the SUBTLEX_{US} database and corresponding rates of production in child speech for relevant low frequency words.

In order to examine whether the limited amount of relevant data available to children could be enough to learn from, I examined individual lexical items which are similarly rare in speech. Only half of the sample words in this frequency range are included on the MCDI; those that are included are reported to be produced by many to most 30 month-olds. Even those that are not seem intuitively to be simple enough that they may also be in some children's vocabularies by 30 months. So, if words with similarly low frequencies to the relevant Principle C structures can be learned by 30 months, is it possible that the infrequent Principle C structures could be enough to form a generalization from? We consider next some challenges to this comparison.

One concern to be raised here is that this is a comparison of two different kinds of frequency. In one case, we have the frequency of sentences that exhibit a particular syntactic construction; in the other case, we have the frequency of a lexical item. Generalizing over multiple pronunciations of a lexical item to determine the meaning seems to be a fairly different task than generalizing over a set of sentences to determine which interpretations are and are not available. When considering the type of encoding

and memory processes involved for learning from each of these sets of data, it becomes clear that the learning process for each may be quite different (and thus, comparison between the two a less compelling argument).

Another concern with this comparison between low-frequency words and low-frequency backward anaphoric structures is the asymmetry in the type of evidence available to the child in each case. As noted above, learning a word will occur via positive evidence—that is, children hear a word uttered, and thus have evidence that it is a well-formed, grammatical element in the language they are learning. Their task is then to form a mapping between this form they have observed and its meaning. In the case of interpretive constraints, children have to rely on indirect negative evidence—that is, children must recognize that a coreferential interpretation is *never* available for sentences with a particular structure. Children must then infer that such an interpretation never occurs because it is ungrammatical. Learning via indirect negative evidence requires an additional level of inference, and thus could prove to be more challenging or time-consuming for the language learner. Although comparison to the acquisition of low-frequency words would suggest that learning from such limited data could be possible, the differences in what must be learned in each case makes this comparison ultimately ineffectual.

In conclusion, we have shown here some initial estimates for the quantity of data available to children for generating interpretive constraints such as Principle C. The low frequency of relevant input and limited developmental time span during which such learning could take place serve as restrictions on any proposed learning mechanism. While I will leave it to the reader to come to their own conclusion about the possible

innateness of Principle C or the viability of a learning account, it should be clear that the research presented here has significantly narrowed the field in terms of how a learning account would be applied.

Question 5: Do children's early Principle C effects contribute to identifying the nature of Principle C as a primitive of grammar?

Much of the exploration of the acquisition of Principle C has been with the goal of furthering our understanding of the nature of Principle C. Is Principle C a primitive of the grammar as the standard view of Binding Theory suggests? Some researchers have proposed alternative ways of accounting for Principle C effects, suggesting that a formal specification of Principle C is unnecessary because effects are accounted for by existing pragmatic principles. This research presents considerable challenges to these views, suggesting that Principle C effects likely cannot be derived from pragmatic sources. We examine two pragmatic accounts here, and show how this research casts doubt on each.

Reinhart (1983) proposes a modified account of Binding Theory based in the observation that standard Binding Theory restrictions fail to capture cases like (92-93), adapted from Grodzinsky & Reinhart (1993), which seem to allow coreference in Principle C contexts.

(92) A: *Who is that?*

B: ***He*** is Colonel ***Weisskopf***.

(93) *Everyone knows that Oscar is incompetent. Even **he** finally realized that **Oscar** is incompetent.*

Reinhart's account is that the interpretation of pronouns can be derived in two ways: via a bound-variable interpretation, where the pronoun receives its interpretation anaphorically; and via coreference, where syntactic binding does not apply due to a lack of co-indexation, but both indices are interpreted as referring to the same entity. This system, which assumes that constraints on coreference hold only over bound variable interpretations, accounts for the possible coreferential interpretation in cases like (92-93). Reinhart's approach essentially acts as a reduced version of standard Binding Theory; she assumes a condition on interpretation of each type of pronominal element (that anaphors must be bound within their governing category, as with standard Principle A, and that pronouns must be free in their governing category, as with standard Principle B). Under an account where syntactic binding conditions only apply to cases of variable binding, however, a condition on coreference with R-expressions is unnecessary, as bound variable interpretations are only possible when the bound element occurs in the c-command domain of the binder. Thus classic cases of Principle C contexts, such as speaker B's utterance in (90), must only derive from a free variable interpretation. Such an account therefore allows a means by which sentences (92-93) are grammatical on the coreferential interpretation; a standard Binding Theory account predicts these sentences to be ungrammatical, although they seem to be perfectly acceptable in certain contexts. However, it is important to note that coreferential interpretations obviating binding conditions are not permissible just anywhere, even under Reinhart's account; compare (93) above to (94).

(94) *Oscar is sad. *He thinks that Oscar is incompetent.*

Reinhart's analysis derives appropriate restrictions on coreferential interpretations through application of an interpretive rule, stated in (95). With this rule, over-generation of coreferential interpretations are eliminated, in that coreference without variable binding is only licensed in specialized pragmatic contexts, when it conveys different information to the hearer than a bound variable interpretation would. This idea of conveying a particular meaning relevant to the context is exemplified by the discourse excerpt in (96).

(95) Rule I: NP A cannot corefer with NP B if replacing A with C, C a variable

A-bound by B, yields an indistinguishable interpretation.

(96) A: *Is that John?*

B: *It must be. He's wearing John's coat.* (Conroy et al., 2009:17)

Here, a coreferential interpretation of speaker B's utterance is possible, rather than necessitating the bound variable counterpart *John's wearing his coat*, because it conveys different information. Namely, it is a remark about a particular salient but unidentified individual wearing John's coat, which is distinct from a remark about a known individual (John) wearing their own coat, as conveyed by the bound variable counterpart.

Reinhart's approach suggests an alternate interpretation of the results observed by LCL as well as ours, in which children arrive at a non-reflexive interpretation by

pragmatic inference of the sentence used to describe the scene. Consider again sentences like (97) and (98).

(97) She's patting herself!

(98) She's patting Katie!

By Reinhart's analysis, (98) must derive from a free variable interpretation, because her analysis does not specify binding conditions for R-expressions. Because a reflexive interpretation where a girl pats herself could be generated from a bound variable utterance like (97), Rule I can be used to infer that a sentence like (98) must therefore have an interpretation distinct from the reflexive interpretation possible by (97).

Is this a plausible interpretation strategy for children to utilize? Beyond recognition of the sentence structure and potential interpretations, children must also employ some hefty pragmatic inference. When children hear a sentence such as (98), they need to be able to understand several important elements. First, they need to recognize that (under Reinhart's analysis) there are two syntactically licensed interpretations available for (98): a coreferential interpretation and a disjoint interpretation. Next, children need to identify a case where replacing the object NP of (98) yields a bound variable interpretation. In this case, this requires replacing the R-expression *Katie* with the reflexive *herself*, as in (97). The way Rule I is specified requires recognizing the interpretations of both of these sentences, even though only one appears in the discourse. Additionally, children must be aware of the discourse context in which the sentence (98) occurs in order to verify that the coreferential interpretation of (98) is indistinguishable

from the bound variable interpretation of (97). Indeed, in the minimal context provided in the Principle C tasks presented here, there is no special context which would create a difference between a coreferential interpretation of (98) and (97). In essence, children need to be able to understand and compare the interpretations of both (97) and (98). Then, children must be able to recognize that the speaker could have uttered (97) if a coreferential interpretation was intended, but they did not. From this, children must then infer that because the speaker chose not to utter (97) and give a bound variable (and coreferential) interpretation, then a coreferential interpretation must not be intended.

One concern for this account of children's interpretations is that the application of Rule I requires these complex pragmatic inferences. In order to compute the intended interpretation of (98), children need to be able to compute Gricean inference to determine that the interpretation must be different from that of (97). The Gricean Maxim of Quantity requires a speaker to be as informative as possible. When a speaker utters (98), the listener uses the Maxim of Quantity to assume that this statement is as informative as possible. Given that if a coreferential interpretation is intended, (97) is a more directly informative utterance than (98), the listener uses the Maxim of Quantity to infer that (97) must not be true, and thus a coreferential interpretation was not intended for (98). Is it likely that children are able to use this Gricean reasoning to make this comparison? Research on children's comprehension of scalar implicatures may suggest otherwise.

Comprehension of scalar implicatures requires a surprisingly similar set of inferences. Scalar implicatures attribute implicit meaning beyond the literal meaning of an utterance, as demonstrated in (99).

(99) Some girls like cheese.

→ It is not the case that all girls like cheese.

The quantifiers *some* and *all* are part of a set of scalar expressions which are ranked in order of informativeness. The scale for quantifying expressions is made up of the expressions in (100).

(100) <all, most, many, some, few>.

When a speaker chooses to use less informative expression on the scale, it can be inferred that a more informative expression would be false. Thus to compute a sentence like (99) a listener must recognize both the interpretation of (99), and the interpretation of a sentence like (101), which includes the more informative expression, *all*.

(101) All girls like cheese.

The listener must then infer that because the speaker chose not to utter (101), that (99) must be the most informative; here arises the implicature that (101) must be false. If you recall, this line of reasoning is very similar to that required by Reinhart's Rule I. In both cases, the listener must (a) identify a sentence which would be more informative but was not uttered; (b) use the Maxim of Quantity to recognize that the speaker would have used the other, more informative sentence if it were possible, but they did not; and (c) infer that the speaker did not intend the meaning of the more informative (un-uttered)

sentence. Thus it seems like the case of scalar implicatures offers a close comparison for the type of inference that children would need to be competent with to be able to compute Reinhart's Rule I. However, evidence from research on children's comprehension of scalar implicatures suggests that children do not successfully complete this type of inference until 5-6 years old (Noveck, 2001; Papafragou & Skordos, to appear). If the inference required to enact Rule I is not available until children are much older, it seems that Reinhart's analysis may not be able to account for the behavior of 30 month-olds.

An alternative pragmatic account has been invoked by researchers in the domain of discourse pragmatics and information structure. Researchers in this domain have taken the preferences for disjoint interpretation in backward anaphoric structures which are not predicted by syntactic restrictions alone as evidence that coreference possibilities might be better accounted for by some other mechanism than syntactic theory. (Note that we have already identified that many of these effects are ameliorated with appropriate discourse context.)

A recent account by Ambridge, Pine & Lieven (in press) argues that "[Principle C] is successful only to the extent that it correlates with principles of discourse and information structure." APL identify the following principles of information structure which are suggested to account for the set of available interpretations.

- (102) "Most utterances have a *topic* (or *theme*) about which some new information (the *focus*, *comment*, or *rheme*) is asserted." These notions can be compared to those of the subject and predicate.

- (103) New topics are expressed with indefinite or full NPs, while topics already established in the discourse are expressed with zero-marking or pronouns.
- (104) Elements such as relative clauses and adjuncts serve to add background information, and are distinct from the central assertion of the utterance.

Together, these principles could provide one way of accounting for the interpretations available for (97) and (98). Let's first consider (97); in this case, the topic (or subject) of the sentence is identified with the pronoun *she*, and the R-expression *Katie* occurs in the assertion made about the topic, i.e. the predicate. According to APL, "it makes pragmatic sense to use a lexical NP... as the topic about which some assertion is made, and a pronoun in a part of the sentence containing information that is secondary to that assertion, but not vice versa." In other words, APL identify a restriction on the intersection of pronouns in topic positions and R-expressions in backgrounded information such as relative clauses or adjuncts. Specifically, this restriction prohibits a single referent from being both topical and backgrounded unless it is new information, in which case the first reference to that entity must be a name and not a pronoun.

This generalization would technically account for Principle C effects in the set of sentences we probed in Experiment 4, repeated here as (105) and (106).

(105) She's painting the house that's in Katie's lap.

(106) The house that she's painting is in Katie's lap.

However, as a generalization to account for all Principle C effects, it fails, as it is limited by being defined over ‘backgrounded information.’ This generalization cannot account for Principle C effects in single clauses, like those tested by LCL or Experiment 4, repeated here as (107) and (108).

(107) She’s patting Katie.

(108) She’s patting herself.

APL utilize a different generalization to account for Principle C effects in single clause utterances: “if a pronoun is used as the topic, this indicates that the referent is highly accessible, rendering anomalous the use of a full NP *anywhere* within the same clause.” So although APL have managed to capture most Principle C effects, their pragmatic account requires two separate generalizations to do so.

We can consider the plausibility of this account given our findings presented here. This research has significantly lowered the age at which Principle C effects are observed, thus lowering the age at which children would need to have command of the pragmatic concepts relevant to implementing APL’s generalizations. If early Principle C effects are to be accounted for by APL’s pragmatic account, then children need to have an adequate understanding of the concepts of topic and focus, as well as having access to the pragmatic principles which govern when pronouns vs. R-expressions are likely to be used. Research suggests that children fail to understand pronouns in the same way as adults far later than our lower bound of 30 months. Jennifer Arnold and colleagues have showed that children fail to use discourse cues to interpret pronouns through age 5;

children do not expect a subject pronoun to refer to the subject of the previous sentence (Arnold et al., 2001; Arnold et al., 2004). Because children seem to still lack some pragmatic cues that would be necessary to implement APL's generalization as late as 5 years old, it seems highly unlikely that they would be able to utilize such a generalization to drive interpretation at 30 months.

In considering pragmatic implementations designed to account for Principle C effects, the results presented here serve to act as a gauge for when children would need to have access to the pragmatic principles in order to attribute children's behavior to such implementations. Independent research shows in both cases we have explored here that children do not have access to the relevant pragmatic components until much later. This suggests that Principle C effects at 30 months (and consequently, any age) should not be attributed to pragmatic effects.

5.2 Conclusion

The research presented in this dissertation has explored the developmental trajectory of Principle C by comparing to other potential interpretive mechanisms as well as by identifying the component processes implicated in interpretation in Principle C contexts. Taken together, the results suggest knowledge of Principle C as the most likely mechanism to drive Principle C effects as young as 30 months.

While the research presented here does not directly contribute to the debate surrounding the nature of such constraints, it does provide evidence about two aspects of the acquisition pattern which could potentially further restrict a learning account of Principle C knowledge. First, our work continues to extend downward the age at which

Principle C effects have successfully been demonstrated to even younger children than the majority of previous research. This narrowing of the age range when Principle C may become active in the child grammar effectively serves to limit the amount of time across development in which the constraint could be learned. Second, this research explores the mechanisms by which children at these youngest ages implement the knowledge responsible for their interpretations, with the aim to diagnose early Principle C behavior as stemming from knowledge of Principle C rather than an alternate interpretive heuristic. Finally, I identified a cursory examination of the input available to children that might allow learning a generalization to take place, showing it to be exceedingly rare. Together, these factors all speak to the feasibility of a learning account, putting us closer on the path to identifying the nature of the Principle C constraint.

The research presented herein serves to provide new insights into methods for exploring early linguistic competence. Because children's competence is not directly observable, we must rely on observation of performance, which is confounded by multiple factors. My research has identified two distinct methods for diagnosing competence. First, comparison of performance across many linguistic contexts can be used to differentiate between multiple potential interpretive mechanisms. Additionally, I introduced a novel approach, which utilizes independent measures of children's processing speed. By creating a proxy for some of the component processes required to implement grammatical knowledge, we are able to compare this to children's observed behavior. In this way, we are able to identify which deployment processes are most relevant to predicting children's performance. Because we know that different knowledge states require different component processes to be implemented, this research allows us

to diagnose underlying competence by examining these deployment processes. Together, these methods of analysis serve to provide new depth to our investigations into children's early syntactic competence. Although we face the challenge of not being able to directly detect underlying competence, these methods of analysis provide ways by which we can infer competence from our careful observation of performance.

Appendix A: Experiment 1 & 2 Stimuli

Lexical Norms for Selected Lexical Items

Lexical Item	Rate of production at 24 months	Rate of production at 30 months
<i>cover</i>	28.1%	60%
<i>dry</i>	37.8%	82.5%
<i>fan</i>	n/a	n/a
<i>paint</i>	31.9%	73.8%
<i>pat</i>	n/a	n/a
<i>spin</i>	n/a	n/a
<i>squeeze</i>	n/a	n/a
<i>wash</i>	63.7%	93.8%
<i>arm</i>	70.4%	100%
<i>body</i>	n/a	n/a
<i>chair</i>	76.3%	95%
<i>face</i>	60.7%	91.3%
<i>head</i>	69.6%	96.3%
<i>shoulders</i>	34.8%	68.8%

Schematic of Task

Phase	Video		Audio
Filler	abstract		<i>classical music</i>
Character Intro	Katie		<i>Oh look! There's Katie. Katie's standing.</i>
Character Intro	Katie		<i>Oh wow! Look at what Katie's doing now. Katie's waving! Look at Katie waving.</i>
Character Intro	Anna		<i>Oh, there's Anna! Anna's standing too.</i>
Character Intro	Anna		<i>Look! Look at what Anna's doing now. Anna's waving! Look at Anna waving.</i>
Character Intro	Anna		<i>Wow, there's Anna again.</i>
Character Intro	Katie		<i>Oh look- there's Katie!</i>
Face Check 1	Anna	Katie	<i>Wow- there they are! Do you see Katie? Where's Katie?</i>
Character Intro	Katie		<i>Yay! There's Katie dancing.</i>
Face Check 2	Anna	Katie	<i>There they are again! Do you see Anna? Where's Anna?</i>
Character Intro	Anna		<i>Yay! There's Anna stretching.</i>
Familiarization 1	Anna dries Anna		<i>Oh wow- there's Katie and Anna! It looks like Anna's body is getting dried!</i>
	Katie dries Anna		<i>Hey look- there they are again! Anna's body is getting dried again!</i>

Test 1	Anna dries Anna	Katie dries Anna	<i>Oh look! Now they're different. She's drying Anna's/her body. Find the one where she's drying Anna's/her body.</i>
Face Check 3	Katie	Anna	<i>Oh look- they're jumping! Do you see Katie? Where's Katie?</i>
Filler	abstract		<i>classical music</i>
Face Check 4	Katie	Anna	<i>Oh look- they're jumping again! Do you see Anna? Where's Anna?</i>
Familiarization 2	Anna pats Katie		<i>Oh wow- there's Anna and Katie! It looks like Katie's head is getting patted!</i>
	Katie pats Katie		<i>Hey look- there they are again! Katie's head is getting patted again!</i>
Test 2	Anna pats Katie	Katie pats Katie	<i>Oh look! Now they're different. She's patting Katie's/her head. Find the one where she's patting Katie's/her head.</i>
Character Intro	Katie		<i>Hey- there's Katie marching! Do you see her marching?</i>
Familiarization 3	Katie paints Anna		<i>Oh wow- there's Katie and Anna! It looks like Anna's face is getting painted!</i>
	Anna paints Anna		<i>Hey look- there they are again! Anna's face is getting painted again!</i>
Test 3	Katie paints Anna	Anna paints Anna	<i>Oh look! Now they're different. She's painting Anna's/her face. Find the one where she's painting Anna's/her face.</i>
Character Intro	Anna		<i>Wow- there's Anna hopping! Do you see her hopping?</i>
Familiarization 4	Anna fans Anna		<i>Oh wow- there's Katie and Anna! It looks like Anna's face is getting fanned!</i>
	Katie fans Anna		<i>Hey look- there they are again! Anna's face is getting fanned again!</i>
Test 4	Anna fans Anna	Katie fans Anna	<i>Oh look! Now they're different. She's fanning Anna's/her face. Find the one where she's fanning Anna's/her face.</i>
Face Check 5	Anna	Katie	<i>Look- there they are again! Do you see Anna? Where's Anna?</i>
Familiarization 5	Anna washes Katie		<i>Oh wow- there's Anna and Katie! It looks like Katie's arm is getting washed!</i>
	Katie washes Katie		<i>Hey look- there they are again! Katie's arm is getting washed again!</i>
Test 5	Anna washes Katie	Katie washes Katie	<i>Oh look! Now they're different. She's washing Katie's/her arm. Find the one where she's washing Katie's/her arm.</i>
Filler	abstract		<i>classical music</i>
Character Intro	Anna		<i>Hey- there's Anna stretching! Do you see her stretching?</i>

Familiarization 6	Anna spins Anna		<i>Oh wow- there's Katie and Anna! It looks like Anna's chair is getting spun!</i>
	Katie spins Anna		<i>Hey look- there they are again! Anna's chair is getting spun again!</i>
Test 6	Anna spins Anna	Katie spins Anna	<i>Oh look! Now they're different. She's spinning Anna's/her chair. Find the one where she's spinning Anna's/her chair.</i>
Character Intro	Katie		<i>Wow- there's Katie dancing! Do you see her dancing?</i>
Familiarization 7	Katie squeezes Katie		<i>Oh wow- there's Anna and Katie! It looks like Katie's shoulders are getting squeezed!</i>
	Anna squeezes Katie		<i>Hey look- there they are again! Katie's shoulders are getting squeezed again!</i>
Test 7	Katie squeezes Katie	Anna squeezes Katie	<i>Oh look! Now they're different. She's squeezing Katie's/her shoulders. Find the one where she's squeezing Katie's/her shoulders.</i>
Face Check 6	Anna	Katie	<i>Look- now they're waving! Do you see Katie? Where's Katie?</i>
Familiarization 8	Anna covers Katie		<i>Oh wow- there's Anna and Katie! It looks like Katie's body is getting covered!</i>
	Katie covers Katie		<i>Hey look- there they are again! Katie's body is getting covered again!</i>
Test 8	Anna covers Katie	Katie covers Katie	<i>Oh look! Now they're different. She's covering Katie's/her body. Find the one where she's covering Katie's/her body.</i>
Filler	abstract		<i>classical music</i>

Appendix B: Experiment 3 Stimuli

Lexical Norms for Selected Lexical Items

Lexical Item	Rate of production at 30 months
<i>cut</i>	80%
<i>feed</i>	81.3%
<i>hug</i>	98.8%
<i>open</i>	96.3%
<i>paint</i>	73.8%
<i>read</i>	92.5%
<i>tickle</i>	87.5%
<i>wash</i>	93.8%
<i>bear</i>	93.8%
<i>book</i>	97.5%
<i>box</i>	96.3%
<i>cat</i>	95%
<i>cake</i>	95%
<i>frog</i>	88.8%
<i>house</i>	92.5%
<i>truck</i>	98.8%

Schematic of Task

Phase	Video		Audio
Filler	abstract		<i>classical music</i>
Character Intro	Katie		<i>Oh look! There's Katie. Katie's standing.</i>
Character Intro	Katie		<i>Oh wow! Look at what Katie's doing now. Katie's waving! Look at Katie waving.</i>
Character Intro	Anna		<i>Oh, there's Anna! Anna's standing too.</i>
Character Intro	Anna		<i>Look! Look at what Anna's doing now. Anna's waving! Look at Anna waving.</i>
Character Intro	Anna		<i>Wow, there's Anna again.</i>
Character Intro	Katie		<i>Oh look- there's Katie!</i>
Face Check 1	Anna	Katie	<i>Wow- there they are! Do you see Katie? Where's Katie?</i>
Character Intro	Katie		<i>Yay! There's Katie dancing.</i>
Face Check 2	Anna	Katie	<i>There they are again! Do you see Anna? Where's Anna?</i>
Character Intro	Anna		<i>Yay! There's Anna stretching.</i>

Familiarization 1	Anna opens box		<i>Oh wow- there's Katie and Anna! Anna's opening a box!</i>
	Katie opens box		<i>Hey look- there they are again! Now Katie's opening the box!</i>
Test 1	Katie opens box	Anna opens box	<i>What's happening now? The box that she's opening is in Katie's lap. Look! The box that she's opening is in Katie's lap. / She's opening the box that's in Katie's lap. Look! She's opening the box that's in Katie's lap.</i>
Face Check 3	Katie	Anna	<i>Oh look- they're jumping! Do you see Katie? Where's Katie?</i>
Filler	abstract		<i>classical music</i>
Face Check 4	Katie	Anna	<i>Oh look- they're jumping again! Do you see Anna? Where's Anna?</i>
Familiarization 2	Katie hugs frog		<i>Oh wow- there's Anna and Katie! Katie's hugging a frog!</i>
	Anna hugs frog		<i>Hey look- there they are again! Now Anna's hugging the frog!</i>
Test 2	Katie hugs frog	Anna hugs frog	<i>What's happening now? The frog that she's hugging is in Anna's lap. Look! The frog that she's hugging is in Anna's lap. / She's hugging the frog that's in Anna's lap. Look! She's hugging the frog that's in Anna's lap.</i>
Character Intro	Katie		<i>Hey- there's Katie marching! Do you see her marching?</i>
Familiarization 3	Anna paints house		<i>Oh wow- there's Katie and Anna! Anna's painting a house!</i>
	Katie paints house		<i>Hey look- there they are again! Now Katie's painting the house!</i>
Test 3	Anna paints house	Katie paints house	<i>What's happening now? The house that she's painting is in Katie's lap. Look! The house that she's painting is in Katie's lap. / She's painting the house that's in Katie's lap. Look! She's painting the house that's in Katie's lap.</i>
Character Intro	Anna		<i>Wow- there's Anna hopping! Do you see her hopping?</i>
Familiarization 4	Katie feeds dog		<i>Oh wow- there's Katie and Anna! Katie's feeding a dog!</i>
	Anna feeds dog		<i>Hey look- there they are again! Now Anna's feeding the dog!</i>

Test 4	Anna feeds dog	Katie feeds dog	<i>What's happening now? The dog that she's feeding is in Anna's lap. Look! The dog that she's feeding is in Anna's lap. / She's feeding the dog that's in Anna's lap. Look! She's feeding the dog that's in Anna's lap.</i>
Face Check 5	Anna	Katie	<i>Look- there they are again! Do you see Anna? Where's Anna?</i>
Familiarization 5	Katie cuts cake		<i>Oh wow- there's Anna and Katie! Katie's cutting a cake!</i>
	Anna cuts cake		<i>Hey look- there they are again! Now Anna's cutting the cake!</i>
Test 5	Anna cuts cake	Katie cuts cake	<i>What's happening now? The cake that she's cutting is in Anna's lap. Look! The cake that she's cutting is in Anna's lap. / She's cutting the cake that's in Anna's lap. Look! She's cutting the cake that's in Anna's lap.</i>
Filler	abstract		<i>classical music</i>
Character Intro	Anna		<i>Hey- there's Anna stretching! Do you see her stretching?</i>
Familiarization 6	Anna reads book		<i>Oh wow- there's Katie and Anna! Anna's reading a book!</i>
	Katie reads book		<i>Hey look- there they are again! Now Katie's reading the book!</i>
Test 6	Anna reads book	Katie reads book	<i>What's happening now? The book that she's reading is in Katie's lap. Look! The book that she's reading is in Katie's lap. / She's reading the book that's in Katie's lap. Look! She's reading the book that's in Katie's lap.</i>
Character Intro	Katie		<i>Wow- there's Katie dancing! Do you see her dancing?</i>
Familiarization 7	Anna tickles bear		<i>Oh wow- there's Anna and Katie! Anna's tickling a bear!</i>
	Katie tickles bear		<i>Hey look- there they are again! Now Katie's tickling the bear!</i>
Test 7	Katie tickles bear	Anna tickles bear	<i>What's happening now? The bear that she's tickling is in Katie's lap. Look! The bear that she's tickling is in Katie's lap. / She's tickling the bear that's in Katie's lap. Look! She's tickling the bear that's in Katie's lap.</i>
Face Check 6	Anna	Katie	<i>Look- now they're waving! Do you see Katie? Where's Katie?</i>
Familiarization 8	Katie washes truck		<i>Oh wow- there's Anna and Katie! Katie's washing a truck!</i>
	Anna washes truck		<i>Hey look- there they are again! Now Anna's washing the truck!</i>

Test 8	Katie washes truck	Anna washes truck	<i>What's happening now? The truck that she's washing is in Anna's lap. Look! The truck that she's washing is in Anna's lap. / She's washing the truck that's in Anna's lap. Look! She's washing the truck that's in Anna's lap.</i>
Filler	abstract		<i>classical music</i>

Appendix C: Experiment 4 Stimuli

Lexical Norms for Selected Lexical Items

Lexical Item	Rate of production at 30 months
<i>ball</i>	100%
<i>bear</i>	93.8%
<i>bike</i>	97.5%
<i>bird</i>	98.8%
<i>block</i>	92.5%
<i>boat</i>	92.5%
<i>book</i>	97.5%
<i>bowl</i>	91.3%
<i>box</i>	96.3%
<i>bus</i>	96.3%
<i>car</i>	98.8%
<i>cat</i>	95%
<i>chair</i>	76.3%
<i>cookie</i>	100%
<i>cup</i>	97.5%
<i>dog</i>	97.5%
<i>doll</i>	90%
<i>fish</i>	88.8%
<i>flower</i>	96.3%
<i>hand</i>	96.3%
<i>hat</i>	93.8%
<i>horse</i>	97.5%
<i>house</i>	92.5%
<i>keys</i>	42.5%
<i>plate</i>	90%
<i>shirt</i>	93.8%
<i>shoe</i>	98.8%
<i>spoon</i>	98.8%
<i>train</i>	97.5%
<i>truck</i>	98.8%

Schematic of Lexical Access Speed Task

Trial	Video		Audio
1	bird	train	<i>Where's the bird?</i> <i>See the bird?</i>
2	spoon	keys	<i>Where are the keys?</i> <i>See the keys?</i>

3	book	cup	<i>Where's the cup?</i> <i>See the cup?</i>
4	shoe	hat	<i>Where's the shoe?</i> <i>See the hat?</i>
5	ball	flower	<i>Where's the flower?</i> <i>See the flower?</i>
6	dog	cat	<i>Where's the dog?</i> <i>See the dog?</i>
7	horse	chair	<i>Where's the horse?</i> <i>See the chair?</i>
8	cookie	fish	<i>Where's the fish?</i> <i>See the fish?</i>

Schematic of Phrase Structure Integration Speed Task

Trial	Video		Audio
Filler	abstract		classical music
1	large green		<i>Hey look- there are some shirts!</i>
	small red	medium red	<i>Where's the biggest shirt?</i>
2	medium yellow		<i>Oh- do you see the boats?</i>
	small yellow	large blue	<i>Where's the biggest boat?</i>
3	small yellow		<i>Oh hey- look at those cats!</i>
	medium yellow	large blue	<i>Where's the biggest yellow cat?</i>
4	medium red		<i>Hey look- there are some chairs!</i>
	small red	large green	<i>Where's the biggest red chair?</i>
5	small yellow		<i>Oh wow- look at those hands!</i>
	medium yellow	large green	<i>Where's the biggest hand?</i>
6	medium red		<i>Now there are some bikes!</i>
	large blue	small red	<i>Where's the biggest red bike?</i>
7	small green		<i>Oh hey- look at those books!</i>
	large blue	medium green	<i>Where's the biggest book?</i>
8	large red		<i>Wow- look at those hats!</i>
	medium yellow	small yellow	<i>Where's the biggest hat?</i>
Filler	abstract		classical music
9	small green		<i>Now there are some shoes!</i>
	large red	medium green	<i>Where's the biggest green shoe?</i>
10	large yellow		<i>Oh wow- look at those houses!</i>
	medium green	small green	<i>Where's the biggest house?</i>
11	small blue		<i>Hey- look at those cars!</i>
	large green	medium blue	<i>Where's the biggest car?</i>
12	medium yellow		<i>Wow- look at those boxes!</i>
	large red	small yellow	<i>Where's the biggest yellow box?</i>
13	large yellow		<i>Now there are some cups!</i>
	medium green	small green	<i>Where's the biggest green cup?</i>

14	medium red		<i>Hey look- do you see those trucks?</i> <i>Where's the biggest truck?</i>
	small red	large blue	
15	large green		<i>Oh hey- look at the bears!</i> <i>Where's the biggest yellow bear?</i>
	medium yellow	small yellow	
16	small blue		<i>Hey look- do you see those dogs?</i> <i>Where's the biggest blue dog?</i>
	large yellow	medium blue	
Filler	abstract		classical music
17	medium blue		<i>Oh wow- look at those buses!</i> <i>Where's the biggest bus?</i>
	large red	small blue	
18	large green		<i>Oh hey- look at the plates!</i> <i>Where's the biggest blue plate?</i>
	small blue	medium blue	
19	small red		<i>Oh- look at those trains!</i> <i>Where's the biggest red train?</i>
	medium red	large yellow	
20	small blue		<i>Wow- now there are some balls!</i> <i>Where's the biggest ball?</i>
	medium blue	large yellow	
21	medium blue		<i>Hey look- do you see the horses?</i> <i>Where's the biggest blue horse?</i>
	small blue	large red	
22	medium green		<i>Oh- look at those dolls!</i> <i>Where's the biggest doll?</i>
	large red	small green	
23	large yellow		<i>Wow- now there are some blocks!</i> <i>Where's the biggest block?</i>
	small red	medium red	
24	large blue		<i>Oh wow- look at those bowls!</i> <i>Where's the biggest green bowl?</i>
	small green	medium green	
Filler	abstract		classical music

Appendix D: Findings Relating Vocabulary and Processing Speed

Study	Method/ design	Age (months)	Significant relation?
Swingley, Pinto & Fernald (1999)	Word recognition	24	No
Swingley & Aslin (2000)	Word recognition	18-23	No
Fernald, Swingley & Pinto (2001)	Word recognition (whole/ partial word)	18	Yes* (whole words only)
		21	
Fernald (2002)	unknown	26	Yes* (data not published)
Zangl et al. (2005)	Word recognition (normal/ degraded speech)	12	Yes
		24	Yes
Fernald, Perfors & Marchman (2006)	Word recognition	15	Yes
		18	No
		21	No
		25	Yes
Hurtado, Marchman & Fernald (2007)	Word recognition (Spanish monolinguals)	15-37	No* (no <i>unique</i> variance) ²⁶
Hurtado, Marchman & Fernald (2008)	Word recognition (Spanish monolinguals)	18	No
		24	Yes
Marchman, Fernald & Hurtado (2009)	Word recognition (Spanish-English bilinguals)	29-34	Yes* (within language only)

²⁶ Hurtado, Marchman & Fernald (2007) found that while age and vocabulary together accounted for a significant portion of the variance in children's reaction times, vocabulary did not contribute any unique variance in addition. This may stem from the extremely large age range across which data was collected.

References

- Ambridge, B., Pine, J.M., & Lieven, E.V.M. (in press). Child language acquisition: Why Universal Grammar doesn't help. To appear in *Language*
- Arnold, J.E., Brown-Schmidt, S., & Trueswell, J. (2007). Children's use of gender and order-of-mention during pronoun comprehension. *Language and Cognitive Processes*, 22(4), 527-565.
- Arnold, J.E., Eisenband, J.G., Brown-Schmidt, S., & Trueswell, J.C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, 76(1), B13-B26.
- Arnold, J.E., Novick, J.M., Brown-Schmidt, S., & Trueswell, J. (2001). Children's on-line use of gender and order-of-mention for pronoun comprehension. In *Proceedings of the 25th annual Boston University conference on language development* (pp. 59-69).
- Arunachalam, S., Escovar, E., Hansen, M.A., & Waxman, S.R. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes*, 28(4), 417-425.
- Arunachalam, S. & Waxman, S.R. (2010). Meaning from syntax: Evidence from 2 year-olds. *Cognition*, 114, 442-446.
- Baier, R., Idsardi, W., & Lidz, J. (2007). Two-month-olds are sensitive to lip rounding in dynamic and static speech events. *Proceedings of International Conference of Audio-Visual Speech Processing*.
- Baker, M. (1991). On some subject/object non-asymmetries in Mohawk. *Natural Language and Linguistic Theory*, 9, 537-576.
- Baker, M. (2001). The natures of nonconfigurationality. In Mark Baltin and Chris Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (pp. 407-438). Oxford: Blackwell.
- Bates, D., Maechler, M., & Bolker, B. (2013). lme4: Linear mixed-effects models using Eigen and R syntax. R package version 0.999999-2. <http://CRAN.R-project.org/package=lme4>.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Brysbaert, M. & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, 41(4), 977-990.

- Bunger, A. & Lidz, J. (2004). Syntactic Bootstrapping and the Internal Structure of Causative Events. *Proceedings of 28th Boston University Conference on Language Development*. Cascadilla Press: Cambridge.
- Bunger, A. & Lidz, J. (2006). Constrained Flexibility in the Acquisition of Causative Verbs. *Proceedings of the 30th Boston University Conference on Language Development*. Cascadilla Press: Cambridge.
- Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, 27(6), 843-873.
- Charney, R. (1980a). Speech roles and the development of personal pronouns. *Journal of Child Language*, 7, 509-528.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.
- Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2009). Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry*, 40, 446-486.
- Crain, S. (1991). Language learning in the absence of experience. *Behavioral and Brain Sciences*, 4, 597-650.
- Crain, S. & McKee, C. (1985). The acquisition of structural restrictions on anaphora. *Proceedings of North East Linguistic Society (NELS) 15*. Amherst, MA: GLSA.
- Crain, S. & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to research on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.
- Crawley, R.A., & Stevenson, R.J. (1990). Reference in single sentences and in texts. *Journal of Psycholinguistic Research*, 19, 191-210.
- Dale, P.S., Dionne, G., Eley, T.C., & Plomin, R. (2000). Lexical and grammatical development: A behavioural genetic perspective. *Journal of Child Language*, 27, 619-642.
- Dale, P.S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, 28, 125-127.
- Devescovi, A., Caselli, M.C., Marchione, D., Pasqualetti, P., Reilly, J., Bates, E. (2005). A cross-linguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, 32, 759-786.
- Dillon, B. (in press). Syntactic memory in the comprehension of reflexive dependencies: an overview. To appear in *Language and Linguistics Compass*.
- Donaldson, M. & Balfour, G. (1968). Less is more: A study of language comprehension in children. *British Journal of Psychology*, 59, 461-471.

- Ehri, L.C. (1977). Do adjectives and functors interfere as much as nouns in naming pictures? *Child Development*, 48, 697-701.
- Eisele, J. A. (1988). *Meaning and form in children's judgments about language*. Cornell University, Aug.
- Eisele, J., & Lust, B. (1996). Knowledge about Pronouns: A Developmental Study Using a Truth-Value Judgment Task. *Child Development*, 67(6), 3086-3100.
- Eisele, J., & Lust, B. (1996). Knowledge about pronouns: A developmental study using a truth- value judgment task. *Child Development*, 67, 3086-3100.
- Fenson, L., Dale, P. S., Reznick, S., Bates, E., Thal, D., Pethick, S. J., et al. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 9(5), v-185.
- Fernald, A. (2002). *How infants develop expectations about what's coming next in speech*. Unpublished paper presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York, NY.
- Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, 42, 98-116.
- Fernald, A., Pinto, J., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, 9, 72-75.
- Fernald, A., Swingley, D., & Pinto, J. (2001). When half a word is enough: Infants can recognize spoken words using partial acoustic-phonetic information. *Child Development*, 72, 1003-1015.
- Fernald, A., Thorpe, K., & Marchman, V.A. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology*, 60.3, 190-217.
- Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. *Language Acquisition and Language Disorders*, 44, 97-135.
- Fiengo, R. (1977). On trace theory. *Linguistic Inquiry*, 35-61.
- Fisher, C., Gleitman, H., & Gleitman, L.R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, 23(3), 331-392.
- Gagliardi, A., Mease, T., & Lidz, J. (submitted). U-shaped development in the acquisition of filler-gap dependencies: Evidence from 15- and 20- month olds. Submitted for publication.

- Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, 124(1), 85-94.
- Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules abstract knowledge of word order in early sentence comprehension. *Psychological Science*, 17(8), 684-691.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1-76.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3-55.
- Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, 14(1), 23-45.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17(3), 311-347.
- Gordon, P.C., & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition*, 62(3), 325-370.
- Gordon, P.C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1411.
- Grimshaw, J., & Rosen, S.T. (1990). Knowledge and obedience: The developmental status of the binding theory. *Linguistic Inquiry*, 187-222.
- Grodzinsky, Y. & Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry*, 24, 69-101.
- Gualmini, A., Hulsey, S., Hacquard, V., & Fox, D. (2008). The question-answer requirement for scope assignment. *Natural Language Semantics*, 16, 205-237.
- Guasti, M. T., & Chierchia, G. (1999/2000). Backward versus forward anaphora: Reconstruction in child grammar. *Language Acquisition*, 8(2), 129-170.
- Hamburger, H. & Crain, S. (1982). Relative acquisition. *Language Development*, 1, 245-274.
- Hart, B. & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H. Brookes Publishing.
- He, A.X., Meng, X., Williams, A., & Lidz, J. (2014). Thematic role assignment in resultative constructions in English- and Mandarin-learning children. *Poster presented at the 13th International Congress for the Study of Child Language*. Amsterdam, The Netherlands.

- He, A.X., Wellwood, A., Lidz, J., & Williams, A. (2014). When participant structure and argument structure do not match: participant structure construction in adults and prelinguistic infants. *Symposium paper presented at the 13th International Congress for the Study of Child Language*. Amsterdam, The Netherlands.
- Hirsh-Pasek, K. & Golinkoff, R.M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.
- Hollich, G. (2005). SuperCoder: A program for coding preferential looking (Version 1.5). [Computer Software]. West Lafayette: Purdue University.
- Hurtado, N., Marchman, V.A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, 34, 227-249.
- Hurtado, N., Marchman, V.A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, 11, F31-F39.
- Jackendoff, R. (1983). *Semantics and cognition*. MIT press.
- Jackendoff, R. (1990). *Semantic structures*. MIT press.
- Kazanina, N., Lau, E.F., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, 56(3), 384-409.
- Kazanina, N., & Phillips, C. (2001). Coreference in child Russian: Distinguishing syntactic and discourse constraints. *Proceedings of the 25th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, 22(6), 860-897.
- Kuhl, P.K., & Meltzoff, A.N. (1982, December). The bimodal perception of speech in infancy. American Association for the Advancement of Science.
- Lakoff, G. (1968). *Pronouns and reference*. Indiana University Linguistics Club, Bloomington, IN.
- Landau, B., & Gleitman, L. (1985). *Language and experience* (Vol. 174). Cambridge, MA: Harvard University Press.
- Langacker, R. (1966). On pronominalization and the chain of command. (pp. 160-186). In D. Reibel & S. Schane (Eds.), *Modern Studies in English: readings in transformational grammar*.
- Lasnik, H. (1976). Remarks on coreference. *Linguistic Analysis*, 2, 1-22.

- Layton, T.L. & Stick, S.L. (1979). Comprehension and production of comparatives and superlatives. *Journal of Child Language*, 6, 511-527.
- Leddon, E., & Lidz, J. (2005). Reconstruction in child language. *Proceedings of the 30th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Lidz, J. (2007). The abstract nature of syntactic representations: Consequences for a theory of learning. In E. Hoff & M Shatz (Eds.), *Handbook of Language Development*. Blackwell: Oxford.
- Lidz, J. & Gagliardi, A. (2010). Morphosyntactic cues impact filler-gap dependency resolution in 20- and 30-month-olds. *Paper presented at the 2010 Boston University Conference on Language Development*.
- Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months. *Cognition*, 89(3), 295-303.
- Lidz, J., White, A.S., & Baier, R. (submitted). The role of incremental parsing in syntactically conditioned word learning. Submitted for publication.
- Lukyanenko, C., Conroy, A., & Lidz, J. (in press). Is she patting Katie? Constraints on pronominal reference in 30-month-olds. To appear in *Language Learning and Development*.
- Lust, B., Eisele, J., & Mazuka, R. (1992). The binding theory module: Evidence from first language acquisition for Principle C. *Language*, 68, 333-358.
- Lust, B., Loveland, K., & Kornet, R. (1980). The development of anaphora in first language: Syntactic and pragmatic constraints. *Linguistic Analysis Seattle, Wash.*, 6(4), 359-391.
- Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: MIT Press.
- MacWhinney, B. (2000). The CHILDES project. *Tools for Analyzing Talk. Part, I*.
- Marchman, V.A. & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, 21, 339-366.
- Marchman, V.A., Fernald, A., & Hurtado, N. (2010). How vocabulary size in two languages relates to efficiency in spoken word recognition by young Spanish-English bilinguals. *Journal of Child Language*, 37, 817-840.

- Mirman, D., Dixon, J.A. & Magnuson, J.S. (2008) Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59, 475-494.
- Musolino, J., & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics*, 44(4), 817-852.
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357-374.
- Naigles, L., Gleitman, H., & Gleitman, L.R. (1993). Children acquire word meaning components from syntactic evidence. *Language and cognition: A developmental perspective*, 5, 104-140.
- Naigles, L.G., & Kako, E.T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child development*, 64(6), 1665-1687.
- Noble, C.H., Rowland, C.F., & Pine, J.M. (2011). Comprehension of argument structure and semantic roles: Evidence from English-learning children and the forced-choice pointing paradigm. *Cognitive science*, 35(5), 963-982.
- Noveck, I. (2001). When children are more logical than adults: Experimental investigations of scalar implicatures. *Cognition*, 78, 165-188.
- Papafragou, A. & Skordos, D. (to appear). Implicature. In J. Lidz, W. Snyder and J. Pater (eds.), *The Oxford Handbook of Developmental Linguistics*. Oxford: Oxford University Press.
- Pearl, L., & Lidz, J. (2009). When domain-general learning fails and when it succeeds: Identifying the contribution of domain specificity. *Language Learning and Development*, 5(4), 235-265.
- Phillips, C. (2004). Linguistics and linking problems. In M. Rice & S. Warren (Eds.), *Developmental Language Disorders: From Phenotypes to Etiologies*. (pp. 241-287). Mahwah, NJ: Lawrence Erlbaum Associates.
- R Development Core Team. (2010). R: A language and environment for statistical computing.
<http://r-project.org/>
- Radford, A., & Galasso, J. (1998). Children's possessive structures: a case study. *Essex research reports in linguistics*, 19, 37-45.
- Regier, T., & Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition*, 93(2), 147-155.
- Reinhart, T.M. (1976). *The syntactic domain of anaphora* (Doctoral dissertation, Massachusetts Institute of Technology).
- Reinhart, T. (1983) *Anaphora and Semantic Interpretation*. London: Croom Helm.

- Roland, D., Mauner, G., O'Meara, C., & Yun, H. (2012). Discourse expectations and relative clause processing. *Journal of Memory and Language*, 66(3), 479-508.
- Ross, J. (1967). Constraints on variables in syntax. (Doctoral dissertation, Massachusetts Institute of Technology).
- Scott, R.M. & Fisher, C. (2009). 2-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes*, 24, 777-803.
- Snedeker, J., & Trueswell, J.C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, 49(3), 238-299.
- Spelke, E. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology*, 15, 626-636.
- Stevenson, R.J., Nelson, A.W., & Stenning, K. (1995). The role of parallelism in strategies of pronoun comprehension. *Language and Speech*, 38(4), 393-418.
- Strayer, J. (1977). *The development of personal references in the language of two-year-olds*. Unpublished doctoral dissertation, Simon Fraser University, Burnaby, British Columbia, Canada.
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48(3), 542-562.
- Swingley, D. & Aslin, R. (2002). Lexical neighborhoods and the word- form representations of 14 month-olds. *Psychological Science*, 13, 480-484.
- Swingley, D. & Fernald, A. (2002). Recognition of words referring to present and absent objects by 24-month-olds. *Journal of Memory and Language*, 46, 39-56.
- Swingley, D., Pinto, J., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition*, 71, 73-108.
- Syrett, K. & Lidz, J. (2011). Competence, performance and the locality of Quantifier Raising: Evidence from 4-year-old children. *Linguistic Inquiry*, 42, 305-337.
- Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language*, 28(01), 127-152.
- Trueswell, J., Sekerina, I., Hill, N., & Logrip, M. (1999). The kindergarten- path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89-134.
- Viau, J., Lidz, J., & Musolino, J. (2010). Priming of abstract logical representations in 4-year-olds. *Language Acquisition*, 17, 26-50.
- Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, 59(1), 67-95.

- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer Publishing Company, Inc.
- Yuan, S., & Fisher, C. (2009). “Really? She Blicked the Baby?” Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. *Psychological Science*, 20(5), 619-626.
- Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, 6, 179–208.